Atmospheric modeling of airborne GHG observations over Europe using a regional transport model: towards quantitative inversions using multiple species

Dissertation

zur Erlangung des akademischen Grades doctor rerum naturalium

(Dr. rer. nat.)

durchgefürt am Max Planck Institute für Biogeochemie Jena

vorgelegt dem Rat der Chemisch-Geowissenchaftlichen Fakultät

der

Friedrich-Schiller-Universität Jena

von Dipl. Phys. Boschetti, Fabio

geboren am 25. September 1986

in Desenzano, Italy

Gutachter:

1. ....

2. ....

Tag der öffentlichen Verteidigung: ....

Abstract

Long-term observations of greenhouse gases are necessary to improve our understanding of sources and sinks of GHGs and their interaction with a changing climate. Such observations are used in combination with inverse atmospheric transport models to estimate surface-atmosphere exchange fluxes. Most of these observations are nowadays collected by ground-based networks of tall towers or satellites in low orbit. However, in the last decade, a new stream of data is gaining momentum: regularly collected airborne data. Airborne data provide an interesting alternative because by collecting observations along the vertical path of an aircraft it is possible to better understand the vertical structure of the atmosphere.

Originally limited by the cost of rental aircrafts, this new source of data can now provide in-situ measurements on a regular basis thanks to strategic partnerships between academia and airlines all over the world. A clever way to reduce costs is in-fact to exploit platforms that are naturally bound to fly as much as possible like commercial airliners. In Europe, the leading project making use of this technique is MOZAIC/IAGOS (Measurements of Ozone by Airbus In-service airCraft / In-service Aircraft for a Global Observing System).

The modeling framework used in this thesis combines a regional Lagrangian transport model (STILT, the Stochastic Time Inverted Lagrangian Transport model) with simulated fluxes from anthropogenic emissions for three trace gases (CO_2, CO and CH_4), biogenic emissions for CO_2, and lateral boundary conditions from global models. We chose this framework because it ensures a fairly good representation of trace gas distribution, it allows for inverse modeling to retrieve regional fluxes, and is flexible enough to assess sources of mismatch between simulated and observed trace gas distributions.

Utilizing a fleet of commercial airliners, MOZAIC/IAGOS provides atmospheric composition data on a regular basis that are widely used for modeling applications. Due to the specific operational context of the platforms, such observations are collected close to international airports and hence in an environment characterized by high anthropogenic emissions. This provides opportunities for assessing emission inventories of major metropolitan areas around the world, but also challenges in representing the observations in typical chemical transport models. We assess here the contribution of different sources of error to overall model-data mismatch using the example of MOZAIC/IAGOS carbon monoxide (CO) profiles collected over the European regional domain in a time window of five years (2006-2011). The different sources of error addressed in the present study are: 1) mismatch in modeled and observed mixed layer height; 2) bias in emission fluxes and 3) spatial representation error (SRE), related to unresolved spatial variations in emissions.

In the first part of the thesis, the SRE was derived by coupling STILT with CO emission fluxes aggregated to different spatial resolutions. We also use the MACC reanalysis to assess uncertainty related to uncertainty sources of the emission fluxes and spatial representation error. We treat the random and the bias

components of the uncertainties separately and found that 1) and 3) have a comparable impact on the random component for both models, while 2) is far less important. On the other hand, the bias component shows comparable impacts from each source of uncertainty, despite both models being affected by a low bias of a factor of 2-2.5 in the emission fluxes for 2006-2011 In addition, we suggested methods to correct for biases in emission fluxes and in mixing heights. Lastly, the evaluation of the spatial representation error against model-data mismatch between MOZAIC/IAGOS observations and the MACC reanalysis revealed that the SRE accounts for roughly 15-20% of the model-data mismatch uncertainty.

In the near future, the scope of the IAGOS program will be expanded by a greenhouse gas measurements system, which allows for the simultaneous measurement of CO_2, CO, and CH_4. Such observation profiles from take-off and landing in the vicinity of major metropolitan areas will be used a.o. as a constraint for sources and sinks. A proposed improvement of the top-down method to constrain sources and sinks is the use of a multispecies inversion. In fact, different species such as CO_2 and CO have partial overlapping in emission patterns for given fuel-combustion related sectors. This implies correlations in the uncertainties, both related to the a priori knowledge of emissions, and to model-data mismatch error. Within the frame of this thesis, we evaluate the benefits of explicitly considering the correlations between uncertainties of different species on the performance of the atmospheric inversion scheme. The available IAGOS CO observations are also used in order to validate the modeling framework.

In the second part of the thesis, we apply the modeling framework to synthetic IAGOS profile observations to evaluate the benefits using correlation between different species' uncertainties on the performance of the atmospheric inversion. The available IAGOS CO observations are used to validate the modeling framework. We found that our modeling framework reproduces the CO observations with an average correlation of 0.78, but simulates lower mixing ratios by a factor 3.8 reflecting a low bias in the emission inventory for 2011. Mean uncertainty reduction achieved for CO_2 fossil fuel emissions is roughly 37%; for photosynthesis and respiration flux it is 41% and 45%, respectively. For CO and CH_4 the uncertainty reduction is roughly 63% and 67% respectively. Considering correlation between different species, posterior uncertainty can be reduced by up to 23%; such reduction depends on the assumed error structure of the prior and on the considered timeframe. The study suggests a significant constraint on regional emissions using multi-species inversions of IAGOS in-situ observations.

Due to delays in achieving the necessary aeronautic certifications for operating onboard a commercial airliner, the plan involving the GHG measurement system had been modified, and the sensor is currently flying on a different research platform: the HALO (High Altitude LOng range) aircraft. HALO is equipped with a rich panoply of sensors, which includes an experimental lidar called CHARM-F capable of parallel retrievals of CO_2 and CH_4 partial columns. The goal is to make CHARM-F traceable to WMO metrological standards using in-situ measurements collected using the JIG sensor during the CHARM-F validation campaign. JIG (Jena Instrument for Greenhouse gas measurements) is in fact already traceable to WMO global meteorological standards for CO_2 and CH_4. The validation of CHARM-F represents an important milestone not only for HALO, but also for the MERLIN space mission from ESA, which is to be developed building upon the CHARM-F's technology.

The focus in this third and last part of the thesis is on two of the HALO instruments: the in-situ JIG sensor and the CHARM-F lidar. We set up a receptor-oriented modelling framework to act as a bridge between the two measurement systems; the main idea is to calibrate the model to match the in-situ measurements, and then use the calibrated model to predict the uncertainty of the lidar vertical partial column (VPC). Slanted partial columns (SPC) were derived from in-situ observations along the flight track. The model-data mismatch was then calculated for the SPC, which was in turn used to derive the expected uncertainty of VPC via error propagation.

In this phase the spatial representation error (SRE), defined as the uncertainty arising from the horizontal averaging of individual emission sources up to the emission inventory's resolution, is scaled to match the model-data mismatch. The scaled SRE is defined as Estimated Total Uncertainty (ETU), which is then used as a proxy for the model-data mismatch as a whole. This means that the derived values of ETU for vertical columns are the values of the uncertainty to be expected from the upcoming observed CHARM-F retrievals. ETU ranges between 0.8-5.0 ppm for CO_2 SPC and 7.9-24.9 ppb for CH_4 SPC, with mean ETU being respectively 1.3 ppm for CO_2 and 9.5 ppb for CH_4. For vertical partial columns, the ETU is instead about 1.0-4.7 ppm for carbon dioxide and 7.7-28.4 ppb for methane. The mean ETU for CO_2 VPC is 1.9 ppm, while for CH_4 said value is 11.4 ppb; this corresponds to an uncertainty in the traceability of lidar observations to WMO standards of 0.5 ppm for CO_2 and 3 ppb for CH_4, which is already close to the WMO compatibility targets. Values of ETU for vertical columns were finally compared with mean values of modeled VPC and atmospheric variability of the individual VPC profiles. Such test was performed for CO_2 and CH_4 by considering both the whole vertical column below the flight track, or by excluding the first 2 km above the ground. This last test was performed in an effort to simulate the conditions associated with a homogeneous cloud cover. It was found that ETU is between 27% and 50% lower than atmospheric variability, indicating that using the model as a transfer standard is indeed helpful.

In conclusion, in this thesis we first made an attempt at describing sources of uncertainty in simulating MOZAIC aircraft profiles of CO; the information regarding the representation error were implemented in IAGOS database. In addition, we describe a method to increase the constraint of atmospheric multi-species inversion in a computationally cheap way. This work paves the way to future studies making use of upcoming IAGOS measurements of CO_2, CO and CH_4. Finally, in the frame of the HALO project, we derived a model to estimate the uncertainty of future lidar-retrieved vertical partial column using in-situ observation from the same platform and a modeling framework.

Zusammenfassung

Langzeitmessungen von Treibhausgasen verbessern das Verständnis über Quellen und Senken, sowie deren Einfluss auf die Klimaentwicklung. Aus solchen Langzeitmessungen werden mit Hilfe von inverser atmosphärischer Transportmodellierung die räumliche und zeitliche Verteilung von Austauschflüssen zwischen Erdoberfläche und Atmosphäre bestimmt. Die meisten Messungen werden heutzutage an atmosphärischen Beobachtungsstationen (Messtürmen) sowie auch mit einigen Satelliten mit niedriger Umlaufbahn gesammelt. Im letzten Jahrzehnt hat jedoch auch die mit Flugzeugen vorgenommene Messung an Bedeutung gewonnen. Während des gesamten Auf- und Abstieges der Flugzeuge können Messungen durchgeführt werden. Auf diese Weise ist ein besseres Verständnis der vertikalen Struktur der Atmosphäre möglich.

Die verbreitete Nutzung dieser neuen Datenquelle war entscheidend limitiert durch die verbundenen Kosten für das Anmieten der Flugzeuge. Dank einer strategischen Partnerschaft zwischen Wissenschaft und Luftfahrtunternehmen in der ganzen Welt können jetzt In-Situ-Messungen mit Flugzeugen auf kontinuierlich operationeller Basis durchgeführt werden. Die Nutzung von Passagierflugzeugen entlang bereits bestehenden Flugrouten ist eine intelligente kostensenkende Lösung. Dieses wird im europäischen Projekt MOZAIC/IAGOS (*Measurements of Ozone by Airbus In-service airCraft / In-service Aircraft for a Global Observing System*) verwirklicht.

Die vorliegenden Dissertation verwendet für die Modellierung das regionale Lagrangian Transportmodell STILT (Stochastic Time Inverted Lagrangian Transport model). Dieses Modell wurde mit den aus einem Emissionsinventar stammenden anthropogenen Emissionsflüssen für CO_2, CO und CH_4, sowie den aus einem diagnostischen Modell stammenden biogenen Flüssen und Randbedingungen für CO_2, CO und CH_4 von einem globalen Modell gekoppelt. Das Modelsystem wurde aufgrund seiner guten räumlichen Spurengasverteilung, seiner Möglichkeiten für inverse Modellierung, und der Flexibilität zur Identifizierung und Quantifizierung verschiedener Modellunsicherheiten ausgewählt.

Durch die Nutzung regulärer Linienflugzeuge in MOZAIC/IAGOS stehen Daten der atmosphärischen Zusammensetzung auf kontinuierlich operationeller Basis zur Verfügung. Der größte Anteil der Messungen stammt dabei aus der unmittelbaren Nähe internationaler Flughäfen. Dieser ist aufgrund des urbanen Umgebungscharakters durch einen hohen Anteil anthropogener Emissionen geprägt. Einerseits ermöglicht dies eine Validierung von Emissionsinventaren in den Metropolregionen. Anderseits ergeben sich damit aber auch Herausforderungen an die Repräsentierbarkeit der Messungen mit den bestehenden chemischen Transportmodellen. Die einzelnen Beiträge des Modell-Datenfehlers wurden für Kohlenmonoxid (CO) aus MOZAIC/IAGOS Profilmessdaten (Zeitraum 2006 bis 2011) für Europa bestimmt. Die in dieser Arbeit untersuchten Quellen der Unsicherheiten sind dabei: 1) Unterschiede zwischen beobachteten und modellierten Mischungsschichthöhen; 2)

Unsicherheiten der modellierten Emissionen; und 3) Unsicherheiten in der räumlichen Repräsentation der Emissionen (SRE = spatial *representation error*).

Im ersten Teil der Dissertation wurde der SRE für CO-Emissionen durch die Kopplung von STILT mit in verschiedenen räumlichen Auflösungen aggregierten Flüsse bestimmt. Eine Validierung der Unsicherheiten der modellierten Emissionen und des SRE erfolgte dann unter Verwendung von MACC-Reanalysedaten. Der zufällige Anteil und der Bias-Anteil vorgenannter Unsicherheitsquellen wurden einzeln untersucht. Der zufällige Anteil hatte für 1) und 3) einen ähnlich großen Einfluss, während dieser für 2) deutlich geringer ausfällt. Der Einfluss des Bias-Anteiles ist für jede der drei Unsicherheitsquellen vergleichbar groß, obwohl die modellierten Emissionen für beide Modellen um ein Faktor von 2 - 2.5 (für 2006 – 2011) zu klein sind. Eine Bias-Korrektur für die Emissionen und die Mischungsschichthöhen wurde bestimmt und angewendet. Der Vergleich des SRE mit dem Modell-Daten-Fehler zwischen MOZAIC/IAGOS Messungen und dem MACC-Modell zeigte, dass die Unsicherheiten des SRE bei ungefähr 15-20% des Modell-Daten-Fehlers liegen.

Ein künftiger Schwerpunkt des IAGOS Projektes wird auf der Einsatz von Sensoren sein, die simultane kontinuierliche Messungen von CO, CO_2 und CH_4 ermöglichen. Die so gewonnenen Vertikalprofile vom Starts und Landungen der Linienflugzeuge in Metropolregionen können zur Bestimmung von Quellen und Senken eingesetzt werden. Eine Verbesserung der inversen Transportmodellierung ist die gleichzeitige Verwendung mehrere Spurengase für die Inversionen. Dabei haben Spurengase wie CO_2 und CO sich teilweise überlappende Emissionsmuster für einzelne Brennstoffsektoren. Dies führt zu einer Korrelation in den Unsicherheiten der a-priori Annahmen der Emissionen verschiedener klimarelevanter Spurengase sowie für deren Modell-Datenfehler. Die Dissertation untersucht die Verbesserung, die sich aus der Berücksichtigung derartiger Korrelationen für ein atmosphärisches Inversionssystem ergibt. Dabei ermöglichen verfügbare IAGOS Messungen von CO dieses Modellierungssystem auch zu validieren.

Im zweiten Teil der Dissertation wurde das Modellierungssystem auf synthetischen IAGOS Messprofile angewendet, um die Vorteile von Inversionen mit mehreren Spurengasen näher zu untersuchen. Verfügbare IAGOS CO Messungen dienten dabei als Validierung für das Modellierungssystem. Die durchschnittliche Korrelation zwischen Modelldaten und Messdaten für CO beträgt 0.78. Die Abweichung der modellierter Werte zu a-priori Werten um den Faktor 3.8 bei kleinen Mischungsverhältnissen deutet auf einen anzunehmenden Bias im den a-priori Annahmen des verwendeten Emissionsinventars für 2011. Eine mittlere Reduzierung der Unsicherheiten um 37% wurde für die anthropogenen CO_2-Flüsse erreicht. Für den biogenen Anteil ergab sich eine Reduzierung um 41% für die Photosynthese und um 45% für die Respiration. Die Reduzierung der Unsicherheiten war 63% für CO und für 67% für CH_4. Berücksichtigt man die Korrelationen zwischen den verschiedenen Spezies, wurde die a-posteriori -Unsicherheit um 23% reduziert. Diese Reduzierung hängt von der a-priori Fehlerstruktur und den verwendeten Zeitskalen ab. Die

Dissertation zeigt eine signifikante Verbesserung der Unsicherheiten in den Emissionsflüssen bei Verwendung einer Inversion mit mehreren Spurengasen von IAGOS In-Situ-Messungen.

An Bord des Forschungsflugzeuges HALO (*High Altitude LOng range*) wird das mobile Lidar-System CHARM-F eingesetzt, das die Gewinnung zeitgleicher Retrievals von CO_2 und CH_4 Teilsäule ermöglicht Das CHARM-F Lidar-System soll in einer Flugkampagne mit in-situ Messungen des mit an Bord befindlichem JIG (*Jena Instrument for Greenhouse gas measurements*) auf WMO Standard kalibriert werden. Die Validierung des CHARM-F Systems ist daher ein wichtiger Schritt nicht nur für HALO, sondern auch für ESA-Satellitenmission MERLIN, für die auf der CHARM-F Technologie ein Messsystem entwickelt wird.

Der dritte Teil der Dissertation widmet sich daher Modellierungsaspekten der beiden vorgenannten HALO Sensoren CHARM-F und JIG. Das Modellierungssystem ist Rezeptor-orientiert aufgebaut und ermöglicht damit eine gemeinsame Modellierungsgrundlage beider Systeme. Zuerst wird ausgehend von den In-Situ Messungen das Modellsystem so kalibriert, dass modellierte Mischungsverhältnisse möglichst gut mit den In-Situ-Messungen übereinstimmen. Danach wird das so kalibrierte Modell eingesetzt, um die Unsicherheiten der mit dem Lidar gemessen vertikalen Teilsäulen (*Vertical Partial Columns,* VPC) vorherzusagen. Geneigte Teilsäulen (*Slanted Partial Columns*, SPC) wurden aus den In-Situ-Messungen entlang der Flugstrecke bestimmt. Der Modell-Datenfehler für SPC wurde berechnet, und dann eine Vorhersage zu den zu erwartenden Unsicherheiten der VPC zu treffen.

Der räumliche Repräsentationsfehler (SRE) wird bestimmt durch die notwendige horizontale Mittelung einzelnen Emissionsquellen unterhalb der Auflösung des Emissionsinventories. Der auf den Modell-daten-Fehler skalierte SRE wird als geschätzte Gesamtunsicherheit (*Estimated Total Uncertainty,* ETU) bezeichnet, und kann dann als Proxy für den gesamten Modell-Daten-Fehler angesehen werden. Dies bedeutet, dass die für ETU erhaltenden Werte für die Vertikalsäulen auch als die zu erwartenden Unsicherheiten der zukünftig mit CHARM-F erhaltenden Retrievals anzusehen sind. Die ETU reicht dabei zwischen 0.8-5.0 ppm für CO_2 SPC und 7.9-24.9 ppb für CH_4 SPC, d.h. einem durchschnittlichen ETU von 1.3 ppm bei CO_2 und 9.5 ppb für CH_4. Für die vertikalen Teilsäulen (VPCs) ist stattdessen die ETU 1.0-4.7 ppm bei CO_2 und 7.7-28.4 ppb bei CH_4. Der durchschnittliche Modell-Daten-Fehler ETU der VPC ist 1.9 ppm für CO_2, und 11.4 ppb für CH_4; das entspricht einer Unsicherheit in der Rückführbarkeit der lidar-Messungen auf WMO-Standards von 0.5 ppm für CO_2 und 3 ppb für CH_4, was schon in der Nähe der WMO-Kompatibilitätszielen ist. Abschließend wurden die ETU-Werte für VPC mit der atmosphärischen Variabilität individueller VPC-Profile verglichen. Für beide Spezies erfolgte dies einmal unter Berücksichtigung des gesamten Vertikalprofils und einmal nur unter Auslassen der ersten beiden Höhenkilometer, was der Simulierung einer homogenen Wolkenbedeckung entsprach. Es zeigte sich, dass der SRE zwischen 27% und 50% niedriger ist als die atmosphärische Variabilität, was den Nutzen des Modells als Transfer-Standard zwischen in-situ Messungen und Fernerkundungsmessung belegt.

Diese Dissertation quantifiziert erstmals die verschiedenen Quellen der Unsicherheiten für die Modellierung von CO MOZAIC Profilen. Der so bestimmte Repräsentationsfehler wurde für jedes CO MOZAIC Profil in die IAGOS Datenbank ergänzend mit eingefügt. Auch beschreibt die Dissertation eine effiziente Methode zur Verbesserung der Bestimmung von Austauschflüssen in einer Multi-Spezies Inversion. Damit ebnet diese Arbeit den Weg für zukünftige Analysen von CO, CO_2 und CH_4 Profilmessungen im Rahmen von IAGOS. Entwickelt wurde im Rahmen des HALO-Projektes auch ein Modell für die Quantifizierung der Unsicherheiten zukünftiger Lidar Messungen der vertikalen Teilsäulen unter Verwendung der ebenfalls an Board von HALO durchgeführten In-Situ-Messungen.

Contents

Acronyms

ASCENDS Active Sensing of CO_2 over Nights, Days and Seasons
CAMS Copernicus Atmosphere Monitoring Service
COFFEE CO_2 release and Oxygen uptake from Fossil Fuel Emission Estimate
CONTRAIL Comprehensive Observation Network for Trace Gases
CRDS Cavity Ring-Down Spectroscopy
DLR Deutsche zentrum für Luft- und Raumfahrt
ECMWF European Centre for Medium-Range Weather Forecasts
EDAS Eta Data Assimilation System
EDGAR Emission Database of Global Atmospheric Research
ETU Expected Total Uncertainty
EVI Enhanced Vegetation Index
FT Free Troposphere
GEE Gross Ecosystem Exchange
GHG Green House Gas
HALO High Altitude LOng endurance
IAGOS In-service Aircraft for a Global Observing System
IEA International Energy Agency
IPCC Intergovernmental Panel on Climate Change
IPDA Integrated path Differential Absorption
JIG Jena Instrument for Greenhouse gas measurement
LIDAR LIght Detection And Ranging
LPDM Lagrangian Particle Dispersion Model
LSWI Land Surface Water Index
LULUCF Land Use, Land use Change and Forestry
MDM Model-Data Mismatch
MERLIN Methane Remote Sensing Lidar Mission
ML Mixed Layer
MODIS MODerate resolution Imaging Spectroradiometer
MOZAIC Measurements of Ozone and water vapor by in-service AIrbus aircraft
MPI-BGC Max Planck Institute for BioGeoChemistry
NEE Net Ecosystem Exchange
PAR Photosynthetic Available Radiation
PBL Planetary Boundary Layer
SRE Spatial Representation Error
RF Radiative Forcing
RHS Right Hand Side
RMSE Root Mean Square Error
SPC Slanted Partial Column
ST Stratosphere
STILT Stochastic Time Inverted Langrangian Transport model

TCCON Total Carbon Column Observing Network
TM3 Transport Model 3
UTLS Upper Troposphere – Low Stratosphere
VPC Vertical Partial Column
VPRM Vegetation Photosynthesis and Respiration Model
WMO World Meteorological Organization

List of Figures

Figure 2.10: Model-data mismatch correlation matrix (a) used in the multi-species inversion, species correlation matrix S_s (b), temporal correlation matrix S_t (c) and squared measurement uncertainty (d). Note that the measurement uncertainty is expressed in ppm for CO_2 and ppb for CO and CH_4. Each matrix has the same dimensions (1098x1098) reflecting the length of the observation vector, but here only the data of July are plotted to increase visibility. The gray lines indicate different species in the observation vector ('blocks'), ordered as follows from top to bottom and from left to right: total CO_2, CO and CH_4. In the single-species inversion, the correlation value in the off-diagonal 'blocks' of matrix (b) is set to zero. The structure in S_s in (c) is a result of the uneven temporal distribution of the observations within the month.

Figure 2.11: The final rescaling matrix ρ_{prior} (a) and the prior error covariance matrix S_{prior}(b). Note that ρ_{prior} can be defined as the element-wise ratio of S_{prior} and C_{prior}.

Figure 2.12: Emission map from EDGAR with receptor points of the CHARM-F validation campaign (left) and example of STILT-derived footprint for a specific receptor point (right).

Figure 2.13: Mean residual profile for CO_2 on the second day of the CHARM-F validation campaign. The three linear fits and residuals in the Mixed Layer (green), Free Troposphere (black) and Stratosphere (blue) are also plotted. The interpolated correction profile used for the correction is plotted in red.

Figure 3.1: Observed CO mixing ratio for the years 2006-2011 in the lower Troposphere around Frankfurt. The plots show mean monthly values at four different heights. Note that values collected at 1000 m are strongly different than values collected at higher levels. In the spring (March and April) of 2007 and 2008, higher values were collected at 2000 m and above. This is likely due to an unusually high number of spring wildfires in many European countries.

Figure 3.2: Observed MOZAIC/IAGOS profiles of CO near Frankfurt (first panel, on top), together with simulated profiles of CO from STILT/EDGAR coupling and MACC reanalysis (second and third panel respectively). The bottom panel shows the absolute value of the residuals between MOZAIC/IAGOS and MACC profiles.

Figure 3.3: Coefficient of determination (R^2) between modeled and observed CO mixing ratio for both STILT/EDGAR and MACC using profiles collected around Frankfurt's airport in years from 2006 to 2011.

Figure 3.4: Comparison of simulated vs. observation-derived mixing heights for MOZAIC profiles near Frankfurt in 2007. The red line is drawn from the Origin and through the center of mass of the scatter plot, so its slope represents the ratio of the mean simulated and observed value.

Figure 3.5: In the left panel the comparison of simulated vs. observed mixed layer CO enhancements for Frankfurt profiles in 2007 during daytime (10:30-17:30 UTC) is shown. The red line is drawn from the Origin and through the center of mass of the scatter plot, so its slope represents the ratio of the mean simulated and observed value. The right panel shows the correction factors to compensate for a bias in STILT/EDGAR and MACC emission flux (right). Both correction factors and error bars (standard deviations) were derived using a weighted least-squares estimate of the parameters of a non-linear model.

Figure 3.6: Median enhancements of CO for the years 2006-2011 in the mixed layer for Frankfurt (left), London (middle), and Vienna (right), as a function of wind direction. The rightmost x-values indicated "low" represent low wind speeds (<3m/s). Observations are show in blue, STILT/EDGAR simulations in different grey tones (light for coarse, dark for high resolution), and MACC reanalysis results are shown in red. STILT/EDGAR and MACC uncorrected enhancements are shown in green and ochre respectively.

Figure 3.7: As Figure 3.6, but for the standard deviation of the enhancements of CO for the different wind sectors.

Figure 3.8: Realizations of representation error (i.e. differences between STILT simulations at different resolutions, here 10 km and 80 km) for CO plotted against simulated enhancement, and color-coded by season (left) and by airport location (right). Grey lines indicate the 5 and 95%ile of the distribution within 10 bins of simulated enhancement; the yellow line indicates the mean.

Figure 3.9: Random component of the relative representation error for CO for the years 2006-2011 in the mixed layer for Frankfurt (left), London (middle), and Vienna (right), as a function of wind direction. The rightmost x-values indicated "low" represent low wind speeds (<3m/s). STILT/EDGAR simulations are shown in different grey tones (light for coarse, dark for high resolution). Maximum relative error for Vienna at 105 degrees is up to 4.8

Figure 3.10: As Figure 3.9, but for absolute representation error.

Figure 3.11: Random component for representation error of Frankfurt for different wind directions (left) and months (right), plotted against the corresponding model-data mismatch error.

Figure 3.12: Assessment of contribution of different error categories for the city of Frankfurt. The assessment is treated separately for random (upper tab) and bias (lower tab) component. For each component the uncertainty of the correction for mismatch in the mixing height (z_i) and bias in the emission inventories (flux) is shown for both STILT (left) and MACC (center) models; the contribution from the spatial resolution of EDGAR fluxes to STILT/EDGAR uncertainty is shown for each of the considered resolutions (right).

Figure 3.13: Absolute change in CO enhancements due to photochemistry (left) and boundary condition (right) on the whole dataset. Standard deviation of residuals is quantified as 2.6 ppb for photochemistry and 11.1 ppb for the boundary condition. Note the different scale on the x-axis.

Figure 4.1: Mean daily enhancement of mixed layer vs. free tropospheric mole fractions. Modeled mixing ratios are shown as black lines, while the observed CO is shown as blue line. Note that the modeled values for CO have been multiplied by a factor of 3.8, corresponding to the mean ratio between observed and modeled CO enhancements (after z_i correction), to match the observed values.

Figure 4.2: Prior error covariance matrix (left) and corresponding posterior error covariance matrix (right).

Figure 4.3: Prior, posterior and true (pseudo-data) fluxes in physical units aggregated for different fuel types. Note that, as the true fluxes are the result of a random perturbation of the prior, they do not describe an actual situation in the physical world. So, for example, the fact that the true value of CH_4 fluxes in July is lower than the same value in December should not be surprising.

Figure 4.4: Prior, posterior and true (pseudo-data) fluxes in physical units aggregated for different emission sectors. Absolute values of biosphere-atmosphere exchange fluxes of CO_2 are included in (b), but scaled down by a factor of 10. Note that, as the true fluxes are the result of a random perturbation of the prior, they do not describe an actual situation in the physical world. So, for example, the fact that the true value of CO for transport in July is higher than the same value in December should not be surprising.

Figure 4.5: Comparison between prior and posterior monthly uncertainties for the five tracers. The posterior uncertainty is plotted for both the multi-species inversion, accounting for inter-species correlations, and the single-species inversion, in which all of the species are independent.

Both prior and posterior uncertainty are expressed in physical units. The spike in the prior methane uncertainty estimate for the month of March depends on the emission inventory and is related to the cycle of agricultural activities.

Figure 4.6: Benefit of a multi-species inversion over the corresponding single-species (dotted line) per different species per months of the year. The benefit has been tested for the three different cases of Table 2.3. Note that CO_2 refers to fossil fuel emissions only, and RESP and GEE refers to the biospheric fluxes.

Figure 4.7: Benefit of a multi-species inversion over the corresponding single-species (dotted line) per different species and month. The benefit has been tested for a "normal" inversion featuring both prior and model-data mismatch correlation between different species (black) or only one of these two components (red and orange). Results refer to Case 1 of Table 3 (black line of Fig. 4.6). Values derived from Palmer (2006) for the month of March are indicated with a diamond.

Figure 5.1: Comparison for CO_2 (top) and CH_4 (bottom) between mean observation (OBS, black dots), modeled profiles (STILT, blue dots) and advected boundary condition (MACC and TM3, green dots). The mean is calculated over all of the 17 profiles of the CHARM-F validation campaign. Modeled profiles are before (left) and after (right) bias correction.

Figure 5.2: Tabs (a) and (c) show values for slanted partial column (SPC) model-data mismatch and expected total uncertainty (ETU) at different heights, for both CO_2 and CH_4. The values of both model-data mismatch and representation error have been derived using all profiles ("all") or by leaving out profiles in Parma for CH_4 and both Parma and Krakow for CO_2 ("subset"). The box-whiskers plots in tabs (b) and (d) show the distribution of individual values of the ratio between model-data mismatch (MDM) and ETU for the "subset" case. The MDM/ETU values are binned every 1000 m and the number of records in each box is reported on the right.

Figure 5.3: Representation error values at different heights for CO_2 vertical partial columns considering only fossil fuels emissions (a) or biospheric emissions (b). Values have been plotted for all profiles (dashed line) and by removing profiles in Parma and Krakow (solid line), characterized by high spatial variability in the mixed layer.

Figure 5.4: Mean VPC profile for CO_2 and CH_4 derived by excluding the high-MDM ones considering the whole vertical column or by excluding the lower 2 km. The values of ETU and atmospheric variability (sd of individual VPC profiles) are plotted in green and cyan respectively.

List of Tables

Table 2.1: Specific emission sectors accounted for in the state vector and aggregated categories as used in Fig. 4.3.

Table 2.2: Specific fuel types accounted for in the state vector and aggregated categories as used in Fig. 4.4.
(*) The category "Other" is derived by summing the contribution from those processes in which is difficult to establish the specific fuel responsible for the emissions.

Table 2.3: relative uncertainty of the prior fluxes aggregated domain-wide and annual for the different cases

Table 4.1: Overall bias for different species between the prior and both posterior and perturbed prior. The percentage values in parenthesis refer to the corresponding Prior-Truth bias.

Table 5.1: Scaling factor and offset values used to train the representation error derived from all profiles and by considering only a subset without those with unrepresentative values of MDM.

Chapter 1 : Introduction

1.1 Climate change and the role of human activities

According to Plank's law, each body emits electromagnetic radiation in a range of wavelength depending on their surface temperature. In the case of the sun, the temperature at the photosphere corresponds to ultra-violet and visible radiation. Earth's atmosphere is largely transparent to these frequencies, allowing most of the energy from the sun to reach our planet's surface and be absorbed by it. As Earth's surface is far colder than the sun's photosphere, it then re-emits the absorbed radiation with far lower frequencies, corresponding to the infrared band. While Earth atmosphere is almost transparent for ultra-violet and visible light, it is instead opaque for infrared, leading to a delay in the escape of the reflected energy and provoking the atmosphere itself to warm up. This is what we call the *greenhouse effect*. Without the greenhouse effect, the surface temperature of Earth would be -18°C, rather than 15°C, and it wouldn't be suitable for life. However, the Greenhouse Effect is largely dependent on the so-called GreenHouse Gases (GHG) like water vapor (H_2O), but also carbon dioxide, (CO_2), methane (CH_4) and nitrous oxide (N_2O), which represent only a minuscule fraction of the atmosphere. This means that the energy balance of planet and therefore its temperature is relatively easy to alter. Accordingly to the overwhelming majority of the scientific consensus, this is exactly what has been happening in the last century and a half, after the beginning of the industrial revolution.

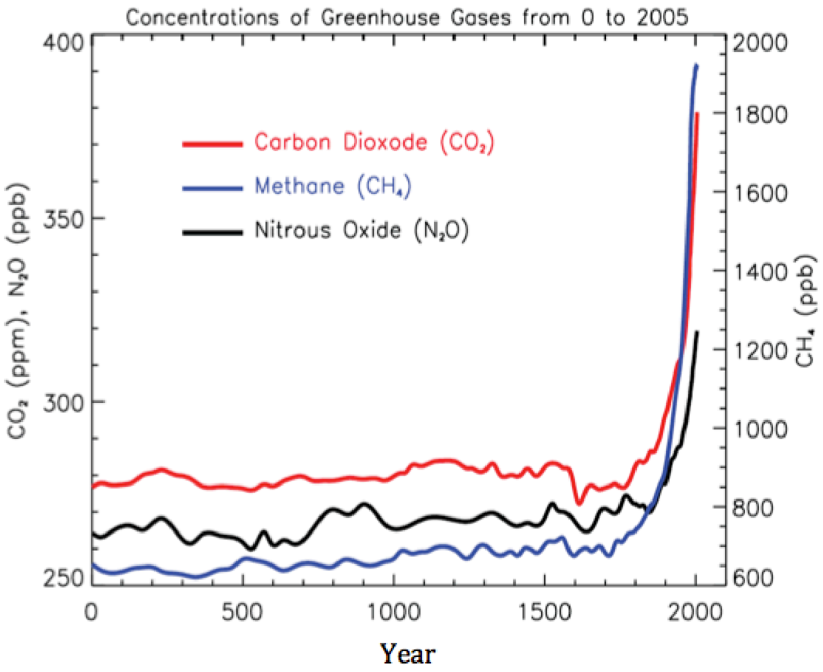

Figure 1.1: Ice-core derived atmospheric concentrations of important long-lived greenhouse gases over the last 2000 years (Forster et al. 2007).

Empirical observation of temperature all around the world together with measured trends of the concentration of the GHG (Fig. 1.1) and theoretical knowledge at elemental particle level has lead the Intergovernmental Panel on Climate Change (IPCC) to declare, in the year 2013, that: *"Warming of the climate system is unequivocal … The atmosphere and ocean have warmed, the amounts of snow and ice have diminished, sea level has risen, and the concentrations of greenhouse gases have increased… It is extremely likely that human influence has been the dominant cause of the observed warming since the mid-20th century".* While it is true that climate change features also natural causes, these are dwarfed by the anthropogenic impact. In addition, despite the fact that some components of the anthropogenic impact are not related to GHG, and are actually characterized by a cooling effect, the net radiative forcing in the atmosphere is a warming effect, mainly due to human activities (Fig. 1.2). Main causes for increase in the concentration of GHG in the atmosphere are fossil fuel combustion, deforestation and change in land use. The two GHGs that are mostly related with human induced climate change are carbon dioxide and methane; global mean CO2 concentration increased from 280 ppm up to the psychological threshold of 400 ppm in 2016 while CH4 increased from 700 ppb to about 1834 ppb (IPCC, 2014).

As human population is expected to increase in the next decades, so it is an increase of human activities and correlated GHG emissions; this makes the need for better policies to mitigate the climate change a matter of greatest importance. One of the major challenges in this respect is to correctly estimate the magnitude of sources and sinks of GHG in a certain spatial domain. One of the most important techniques to solve this problem is the so-called inverse modeling, or "top-down" approach. Inverse modeling makes use of the fact that atmospheric GHG mixing ratios in different locations and times are related to their sources and sinks. Thus it is possible to infer the spatial distribution of surface-atmosphere fluxes from atmospheric observations.

As policymakers engage in talks to curb emissions, it is part of the duty of the scientific community to keep increasing the understanding of climate change through better measurements and to provide more robust and reliable predictions, which need better understanding of these sources and sinks, which in turn requires a combination of dense and long-term observations with improved atmospheric tracer transport models for inverse modeling.

At present, the vast majority of atmospheric observations are collected by either land-based networks of tall towers, or satellites in low orbit, from below and above the atmosphere respectively. The satellites, in particular, make primary use of passive sensors. In the last decade two new streams of data are showing potential to change the landscape in the atmospheric observations community: the first is mixing ratio measurements from airborne platforms, the second is the use of retrievals from active sensors (lidars) mounted on both satellites and aircrafts.

In this dissertation we characterize different sources of model error and prepare inverse modeling tools for using airborne measurements from MOZAIC/IAGOS, in addition we attempt to estimate the traceability of the CHARM-F lidar carried by the HALO research aircraft to WMO international standards.

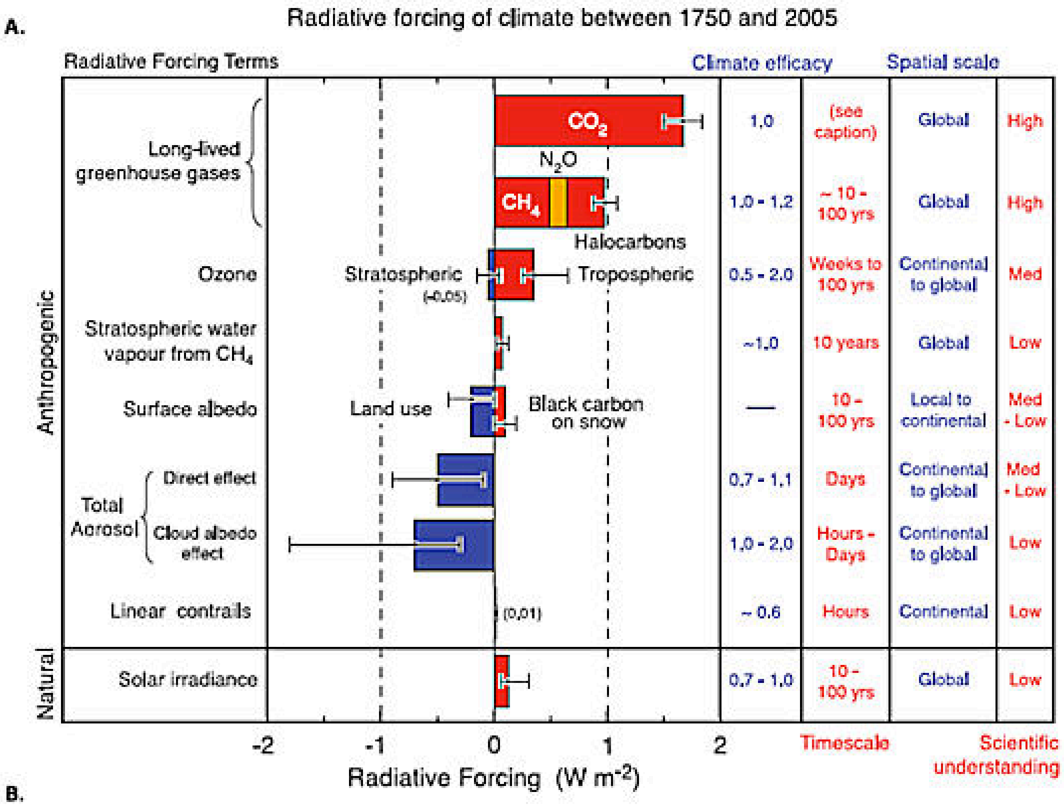

Figure 1.2 Global mean RFs from the agents and mechanisms grouped by agent type. Anthropogenic RFs and the natural direct solar RF are shown. Columns indicate other characteristics of the RF; efficacies are not used to modify the RFs shown. Time scales represent the length of time that a given RF term would persist in the atmosphere after the associated emissions and changes ceased. No CO2 time scale is given, as its removal from the atmosphere involves a range of processes that can span long time scales, and thus cannot be expressed accurately with a narrow range of lifetime values. Figure is obtained from Forster et al. (2007).

1.2 Scientific questions

This thesis is centered on making these two new data streams more usable and is divided into three main parts, each one addressing a different theme. The first part is about describing different sources of error in the modeling of in-situ airborne measurements, the second is about inter-species correlations in the frame of atmospheric inversions, with emphasis on paving the way to future multi-species airborne observations and the third part is focused on validation of an airborne lidar. This thesis addresses a number of scientific questions for each of the main themes:

1. How is model-data mismatch by current transport models distributed? Is the MOZAIC/IAGOS dataset spatially representative at European scale?
2. Is there a benefit in considering the correlations in prior and observed uncertainties of different chemical species in the frame of atmospheric inversions? How large is this benefit? Is it possible to compare atmospheric inversions in which different species are completely uncorrelated with one in which inter-species correlations are considered?
3. How can airborne lidar retrievals of atmospheric columns be validated against airborne in-situ data with the help of a tracer transport model? To what level of uncertainty can this validation of vertical partial column measurements be achieved?

For each part of the thesis, a specific modeling framework was set up to address the scientific questions listed above. Despite their differences, these three modeling frameworks tend to use similar components and are always based on a Lagrangian Particle Dispersion Model (LPDM) and at least one emission inventory. In addition, central for the part on the multi-species inversion, is a Bayesian inversion scheme.

In the first part of Chapter 2 (Sect. 2.1) the single components of the modeling framework are described separately, while in the Sect. 2.2-2.4 the three modeling frameworks, together with related observations and tools for statistical analysis are described in detail. Chapters 3, 4 and 5 describe and discuss the results of the three main themes, while Chapter 6 draws the conclusions.

1.3 Different sources of uncertainty in the modeling of airborne measurements

At present time, the lion’s share of atmospheric observations comes from two main sources: in-situ measurements from ground based observational networks, and remote sensing from satellite-borne instruments. Globally distributed ground based networks measure atmospheric mixing ratios of a number of atmospheric species, including GreenHouse Gases (GHG) such as CO_2 (Rödenbeck et al., 2003) or CH_4 (Hein et al., 1997; Bousquet, et al. 2006), but also chemically active species such as CO (Bergamaschi et al. 2000). Modelers trying to tease apart different sources and sinks in a certain spatial domain often use atmospheric observations from the global network as top-down constraint in inverse modeling. Inverse modeling simulates atmospheric transport using a general circulation model to track different air parcels that are observed. In this way it is possible to deduce magnitude and spatial distribution of sources and sinks in a global domain. As for data from space-borne platforms, the combination of several sensors on different satellites allow for daily global coverage of different species, including the above mentioned CO_2, CH_4 (SCIAMACHY, GOSAT) and CO (MOPITT, AURA) from low orbit. Albeit of lower quality in terms of the measurement uncertainty, due to their coverage also in otherwise inaccessible and sparsely sampled regions those observations have a large potential for inferring e.g. emissions of CH_4 (Bergamaschi et al., 2009), or CO emissions (Kopacz et al., 2009), or sources and sinks of CO_2 (Nassar et al. 2011).

An interesting recent alternative is represented by aircraft measured profiles, which allows for gathering mixing ratio information across the whole vertical path of the flight, leading to a detailed description of the internal structure of the troposphere. Many recent studies made use of aircraft profiles alone or in combination with other data sources (e.g.: Brioude et al., 2013; Gourdji et al., 2012). However, mainly due to the cost of a rental aircraft, the number of flights is usually quite limited, with direct consequences on data availability. A way of overcoming such a limitation is to make use of commercial airliners. This approach makes available atmospheric concentration measurements on a regular basis, and has been selected from research projects such as CONTRAIL (Comprehensive Observation Network for Trace Gases) (Machida et al., 2008), and MOZAIC/IAGOS (Measurements of Ozone and water vapor by in-service AIrbus aircraft / In-service Aircraft for a Global Observing System) (Marenco et al., 1998; Volz-Thomas et al., 2009). MOZAIC/IAGOS has been active for more than two decades by now, and is widely recognized as an important data provider for atmospheric modeling applications and for calibration/validation (Cal/Val) of satellite observations. Among others, observations from MOZAIC/IAGOS (Fig. 1.3) have been used in an attempt to describe a global vertical profiles CO climatology (Zbinden et al.,2013). A more detailed and recent description of the project is available in Petzold et al. (2015).

MOZAIC/IAGOS provides atmospheric composition data collected from long-haul passenger aircraft. This implies that these observations are made in a quite specific context: taking off and landing at major airports, and cruising in flight corridors in the upper troposphere and lower stratosphere. In addition, the observations are made in-situ, as point observations along the flight track. It is obvious that MOZAIC/IAGOS observations are influenced by this specific context, which is characterized by high local anthropogenic emissions. Thus it is possible that MOZAIC/IAGOS observations are representative only at local scale and hence high-resolution models are needed to capture these local features, with direct impact on computational effort. For this reason, understanding the sources of error of such observations is crucial for a successful use of their information content in the context of modeling or Cal/Val.

In the context of modeling, it is paramount to assess how well models can reproduce observations; the difference between model outputs and retrieved measurements (observations) is hereafter referred to as model-data mismatch. Model-data mismatch composes of different error sources, for example:

- Observation uncertainty
- Mixed layer height mismatch
- Uncertainty of the bottom-up derived emission fluxes

- Unresolved spatial variations in emission fluxes

Chapter 3 of the present thesis is focused on a quantitative description of the abovementioned four sources of error related to model imperfection in the frame of the IGAS project. The aim of IGAS (IAGOS for GMES Atmospheric Service) is to improve connections between data collected by MOZAIC/IAGOS and the Copernicus Atmosphere Monitoring Service (CAMS), where the data are used for model evaluation. CAMS is intended to provide continuous data and information on atmospheric composition, both in hindcasting and in forecasting a few days ahead.

The mixed layer is usually defined as the part of the troposphere in which a compound is well mixed due to turbulent convection in the time scale of an hour or less (Seibert et al., 2000), and for this reason it is the part of the troposphere in which surface influence from anthropogenic emissions is strongest. A poor modeling of the vertical mixing transport is well known for being one of the most important sources of error in atmospheric modeling and has been already investigated in at least one recent paper (Kretschmer et al., 2012).

In simulating atmospheric composition, not only transport, but also emission fluxes need to be modeled based on emission inventories. Uncertainty in the simulated fluxes from emission inventories is not a completely new issue. Underestimation of CO mixing ratios by most atmospheric models has led some author to investigate the accuracy of emission inventories (Stein et al., 2014).

Effects from spatial resolution of simulated fluxes can be described in different ways; the quantitative indicator for this source of uncertainty is hereafter referred to as representation error. Assessments of representation error for airborne measurements have been made for several research campaigns by applying spatial statistic methods to densely distributed profiles of CO_2 mixing ratios (Gerbig et al., 2003a; Lin et al., 2004). This empirically derived representation error was shown to be consistent with a model-based analysis, that combines a Lagrangian transport model with spatially resolved surface-atmosphere fluxes at different spatial resolutions (Gerbig et al., 2003b); this indicated that most of the observed spatial variability of trace gases in the mixed layer is explained by spatial variability in surface-atmosphere fluxes. Here we use a similar approach: we combine the STILT (Stochastic Time Inverted Lagrangian Transport) model (Lin et al., 2003) with high-resolution fossil fuel emission inventories, in order to assess the impact of the spatial resolution on simulated mixing ratios within the PBL, and ultimately on the representativeness of profile observations for specific spatial scales. We use standard deviations of differences resulting from different spatial resolution in simulated CO to quantify the spatial representation error. The central question for this third model-derived source of uncertainty is to which degree spatial and temporal variations in the representation error are meaningful to describe and ideally predict corresponding spatial and temporal variations in the model-data mismatch. If successful, the knowledge of such a representation error would allow for a more quantitative comparison between point observations of mixing ratios and corresponding simulations at coarser spatial scales.

Chapter 3 addresses the partitioning of uncertainties for one of the main parts of the MOZAIC/IAGOS observations: vertical profile data collected during take off and landing. The focus of the work is on carbon monoxide (CO), a non-greenhouse gas that is of interest as a tracer for anthropogenic emissions as emission fluxes of CO are mostly collocated with those of CO_2 from fossil fuel combustion.

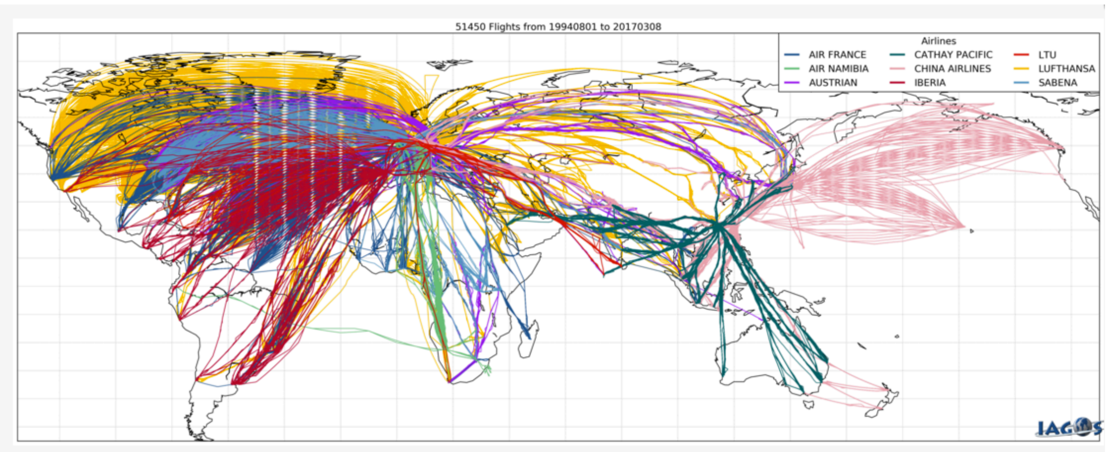

Figure 1.3 Map of MOZAIC/IAGOS flights from 1994 to 2017. Source: http://www.iagos.sedoo.fr.

1.4 Multi-species inversion for future airborne measurements

Climate predictions are currently hampered by excessive uncertainties. A symptom of this is that intercomparisons of different models show important differences in their predictions, as shown in Friedlingstein (2006). This makes it difficult to assess the better environmental policies to implement. As widely recognized at an international level, there is a need for the reduction in anthropogenic emissions (IPCC, 2014). This however implies the necessity for emissions monitoring to verify whether emission-reduction policies are successful. Because most biogenic fluxes in Europe are influenced by human activities, understanding and managing these biogenic fluxes must also be a component of any policy to reduce anthropogenic emissions.

An important tool to estimate carbon budgets by teasing apart sources and sinks in a given spatial domain is the atmospheric Bayesian inversion. Atmospheric inversions combine prior knowledge from emission inventories with atmospheric observations acting as a top-down constraint to produce better posterior knowledge. An important metric to measure the effectiveness of an atmospheric inversion is the uncertainty reduction, defined as the difference between prior and posterior uncertainty normalized with the prior uncertainty. The vast majority of published papers on atmospheric inversions investigate the budget of a single species, usually a long-lived greenhouse gas like CO_2 (e.g. Rödenbeck, 2003) or CH_4 (e.g. Hein, 1997; Bousquet, 2006), but the technique can also be applied to active species like CO (Bergamaschi, 2000). Note that carbon dioxide is a special case as atmospheric CO_2 mixing ratios result from a combination of strong anthropogenic sources with strong sources and sinks from biospheric processes, calling for a separation of anthropogenic from biospheric fluxes. One way to achieve such a separation is to measure CO alongside CO_2, and use CO as a proxy for CO_2 anthropogenic emissions. Palmer (2006) used CO_2-CO correlations to improve an inversion of data from the TRACE-P aircraft mission in March-April 2001, while Wang (2009) employed a similar method using satellite data, obtaining a reduction in CO_2 flux inversion error.

So far the lion's share of the studies investigating atmospheric inversions make use of both continuous in situ and flask measurements from ground based observational networks of tall towers (e.g. Kadygrov, 2015; Sasakawa, 2010). However, as profiles collected from an aircraft easily exceed the height of towers, airborne data may also prove an interesting option for this application. This alternative was tested in some recent studies that made use of aircraft profiles alone or in combination with other data sources (e.g.: Brioude, 2013; Gourdji, 2013). Methods to maximise the cost-effectiveness of airborne data are the use of unmanned aircraft (drones) and commercial airliners. The latter, in particular, allows for collecting data on a regular basis without requiring a particularly small or light sensor.

The most important projects making use of commercial airliners are CONTRAIL (Comprehensive Observation Network for Trace Gases) (Machida, 2008), and MOZAIC/IAGOS (Measurements of Ozone and water vapor by in-service AIrbus aircraft / In-service Aircraft for a Global Observing System) (Marenco, 1998; Petzold, 2015). Both projects have been running for more than two decades and have produced extensive datasets that have proven to be important in the fields of atmospheric modeling and satellite calibration and validation. Regarding carbonaceous species, CONTRAIL has so far been collecting CO_2 mixing ratio measurement, while IAGOS was focused on CO. In the next years the IAGOS fleet will simultaneously provide CO, CO_2 and CH_4 atmospheric concentration measurements (Filges, 2015), enabling the use of multi-species synergy in modeling applications. This synergy follows from the fact that the collocated measurements share the same atmospheric transport and have partially correlated emission uncertainties.

Chapter 4 of this thesis is focused on investigating the benefits on uncertainty reduction of such a multi-species inversion in comparison with a single-species inversion. To attain this goal, we set up a synthetic experiment utilizing the measurement times and locations collected from the IAGOS projects in the year 2011. This study is intended to pave the way for future studies making use of multi-species IAGOS datasets when they become available.

Chapter 4 builds upon the previous one by adding other species, and using a formal Bayesian inversion. A multi-species inversion was carried out in order to exploit the correlations in uncertainties between CO_2, CO, and CH_4, specifically in their respective uncertainties in a priori anthropogenic emissions and in model representation error. The aim of this multi-species inversion is to provide better estimates of anthropogenic emissions, and, in the case of CO_2, to better separate the biospheric from anthropogenic contributions.

1.5 Validation of airborne lidars

In the last decades, the need of frequent observations of the atmosphere over large or inaccessible regions has been increasingly satisfied by data collected from space-borne platforms; satellites in low orbit are in fact capable of virtually global coverage in a timeframe of 2-3 days. Different satellites provide coverage for greenhouse gases (GHG) like CO_2 (OCO2), CH_4 (SCIAMACHY, GOSAT), and CO (MOPITT, AURA); in addition, by mounting different sensors on the same platform, or by grouping satellites in constellations like the NASA A-Train, it is possible to achieve the simultaneous retrieval of a number of parameters and species.

Despite the important advantages provided by satellites, they also come with a major limitation. In fact, most of the current satellites are equipped with passive sensors and their coverage is therefore limited to daylight and clear-sky condition hours. This problem is exacerbated in presence of scattered clouds as many satellites also provide a rather coarse spatial resolution. In addition, the cost of a space mission is very high, and once the satellite has been launched, it is of course impossible to update and maintain the hardware. This problem can be solved by lidars. Lidars are active sensors and therefore they are independent of solar radiation and allow for day and night measurements. In addition they are able to see through narrow cloud gaps thanks to their small footprint and ranging capability. They are also less sensitive to atmospheric aerosol compared to traditional passive optical sensors. Due to their small laser footprint and high sampling rate, lidars also provide high spatial resolution retrievals, so that they can be used to measure very close to the emission source. Observation in these areas has traditionally been difficult as the contrast to the background concentration is relatively high in proximity of the source. This feature, combined with the near all-weather measuring capacity allows for a better monitoring of strong localized emission sources such as power plants and to better estimate their total emissions. This is also important to increase discrimination of different contribution in case of multiple sources in a small area.

One of the upcoming satellites to be equipped with a lidar is the MERLIN (MEthane Remote sensing LIdar missioN) space mission from ESA, scheduled for launch in 2021 and due to remain in service for three years (Pierangelo et. al, 2016; Ehret et al.,2013). Using an IPDA (Integrated Path Differential Absorption) lidar for CH_4, MERLIN is tasked with creating a world-wide map of methane in the atmosphere, with emphasis on identifying sources and sinks.

Regarding the platform, a recent and increasingly important stream of data that avoids the high cost and difficulties in upgrading the payload of satellites is remote sensing using aircrafts. As aircrafts fly closer to Earth's surface than satellites, they also provide better resolution and for this reason are suitable for applications that require higher resolution like site specific farming (G. B. Senay, 1998); in addition they are cheaper and easier to update than satellites. For this reason, remote sensing instruments for satellites are usually first put on aircraft during their development phase; this is also the case for CHARM-F, which is the "airborne simulator" for the MERLIN lidar instrument. Currently, a limited number of studies have been published about airborne lidars specific for CO_2. In this frame, the most prominent examples are three sensor concepts developed in the frame of the ASCENDS (Active Sensing of CO_2 over Nights, Days and Seasons) mission proposal from NASA (Abshire et al., 2014; Spiers et al.,2011; Browell et al., 2008). A lower number of studies have been involving a CH_4 specific lidar (e.g. Riris et al., 2012).

A new addition to a growing fleet of research aircraft fitted with remote sensing equipment is the HALO (High Altitude LOng endurance) aircraft operated by DLR (Deutsche zentrum fur Luft- und Raumfahrt). HALO is equipped with a rich panoply of sensors for investigating, among others: 1) the influence of aerosol on cloud processes, 2) vertical exchange in the Upper Troposphere – Lower Stratosphere (UTLS), 3) long range transport of pollutants.

In the present study, we focus on two of the HALO sensors: the in-situ JIG (Jena Instrument for Greenhouse gas measurement) sensor and the CHARM-F lidar. Also known as IAGOS-core GHG package, JIG was originally intended for the IAGOS (In-service Aircraft for a Global Observing System) project (Filges et al, 2015; Petzold et al., 2015), and allows for collecting in-situ observations of three different greenhouse trace gases (CO_2, CO and CH_4). CHARM-F is a newly developed airborne IPDA lidar that allows for the retrieval of CO_2 and CH_4 partial columns, and still has to be tied to the WMO standards. Our main goal is to validate the CHARM-F lidar by comparing its retrieval with in-situ measurements from JIG. JIG is already traceable to the WMO meteorological standards, therefore it is possible to use it in order to make retrievals from CHARM-F traceable to WMO primary scales for CO_2 and CH4.

Due to the conceptual differences in the atmospheric sampling by the two involved measurement techniques, a direct sensor-to-sensor comparison is difficult. For this reason, we set up a receptor-oriented modelling framework to act as a bridge between the two measurement systems. The general idea is to calibrate the models to match the in-situ measurements, and then use the calibrated modelling framework to make predictions on uncertainty of the lidar retrievals. To carry out the comparison, the model has to be run for different sets of receptor points (starting locations of the backward Lagrangian particle dispersion model) for the different sensors: directly along the flight track for the in-situ measurements, and in a vertical array of locations placed directly below the flight track to simulate the lidar.

The main components of the framework are the STILT (Stochastic Time Inverted Lagrangian Transport) model, a Lagrangian Particle Dispersion Model (LPDM) for the atmospheric transport of the trace gases, the Emission Database for Global Atmospheric Research (EDGAR) to simulate anthropogenic emissions, the Vegetation Photosynthesis and Respiration Model (VPRM) to model biogenic emissions, a global circulation model (TM3) to provide boundary conditions for CH_4 and the MACC reanalysis to provide boundary conditions for CO_2 and CO.

The validation of CHARM-F paves the way to a more regular use of the lidar in the frame of the HALO research missions and will expand the research focus of the platform. In addition, as an instrument with similar specification is being developed for the MERLIN mini-satellite, the validation of CHARM-F is an important step in the development of that project as well.

Chapter 2 : Material and Methods

2.1 Models Overview

To address the research questions of this thesis (cfr. Sect. 1.2), a modeling framework needs to be established for each part of the thesis. Each modeling framework consists of various components, but the most important are the following: a transport model, two different emission models, and a Bayesian inversion scheme. A simplified representation of how these components interact in this thesis is depicted in Fig. 2.1.

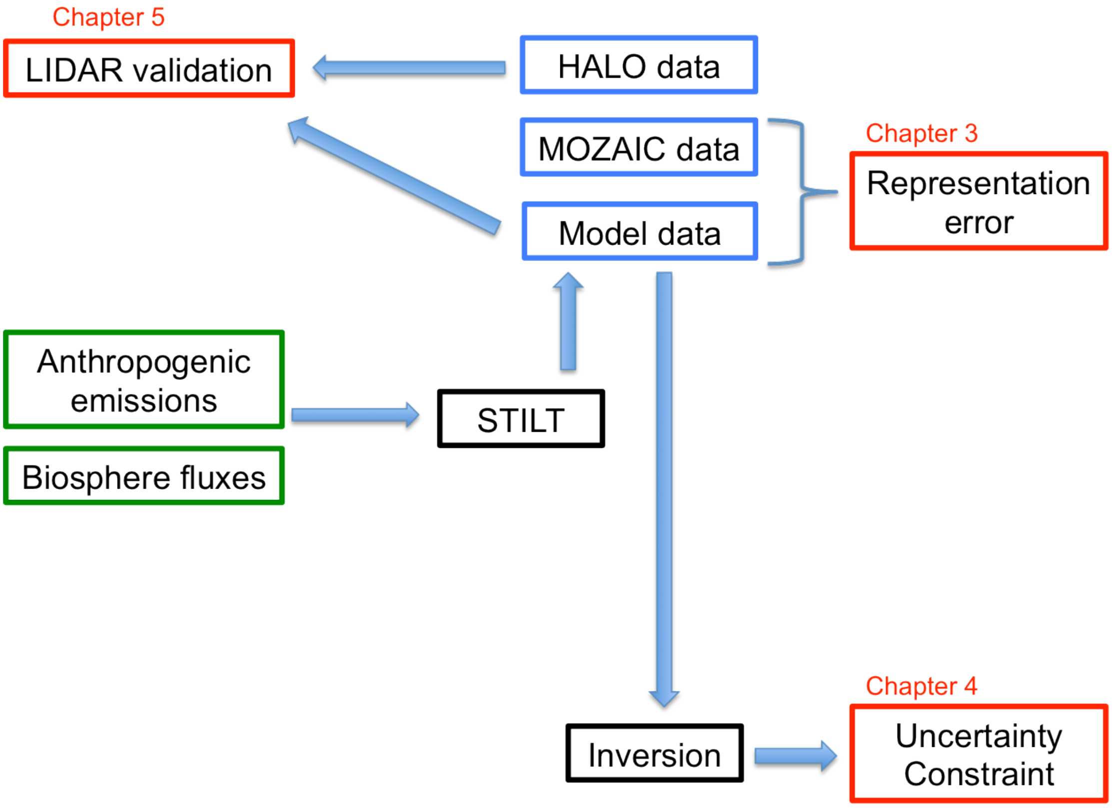

Figure 2.1: Modeling framework scheme

2.1.1 The STILT model

The STILT (Stochastic Time Inverted Lagrangian Transport) model is a Lagrangian Particle Dispersion Model which have been used in regional simulations and inversion studies for different greenhouse gases (e.g. Gourdji et al.,2012; Thomson et al.,2011; Gerbig et al.,2003). It is derived from the source code of the HYSPLIT system (Hybrid Single-Particle Lagrangian Integrated Trajectory) (Draxler and Hess, 1998) but makes the assumption that turbulent diffusion can be modeled as a Markov chain, in which the velocity vector can be decomposed as:

$$u = \overline{u} + u' \tag{2.1}$$

where the right hand side (RHS) of the equation is the sum of a mean component and a turbulent component respectively. A detailed description of the STILT model is given in Lin et al. (2003).

The main purpose of STILT is to calculate *footprints,* defined as the sensitivity of the atmospheric mixing ratio measurements to the upstream surface fluxes. Traditional atmospheric particles dispersion models usually simulate the spread of a cloud of virtual particles from a point location, usually representing pollutants emitted e.g. by a chimney and spread by simulated winds from a meteorological field. This means that to model the atmospheric mixing ratio of a tracer in a given location (receptor), it is necessary to evaluate the tracer's fluxes on the whole domain, as the tracer's origin itself is unknown. As stated by it own name (Stochastic Time-Inverted Lagrangian Transport), a main feature of STILT is the possibility of running the ensemble of particles back in time allowing to track the origin of each particle, and hence identify which portions of the domain actually contributed with their emissions at the tracer's concentration at the receptor and the relative contribution for each of them. In this way it is possible to reduce the required calculation time and consequently increase the framework efficiency.

The virtual particles simulated by STILT represent air parcels of equal mass, transported by wind fields generated by ECMWF (European Center for Medium-range Weather Forecasts). The ensemble of the resulting back-in-time trajectories are used to calculate the footprints, that are later coupled with emission inventories databases and concentration fields at the starting time, in order to simulate the tracer's concentration at the receptor location. In this way, the footprints produced by STILT can be used as the adjoint transport model that is needed for the atmospheric inversion. ECMWF simulations are derived with a spatial resolution of about 10 Km and 3-hourly temporal intervals.

STILT is a mere regional transport model; as such it needs boundary conditions to take into account the effects that the rest of the global circulation has on its regional domain. In addition, the produced footprints can be coupled with different emission inventories to derive simulated mixing ratios of different trace gases. The TM3 global tracer transport model (Heinmann and Körner, 2003) and MACC reanalysis will provide the initial/boundary condition of the European domain, while EDGAR (Emission Database for Global Atmospheric Research) and VPRM (Vegetation Photosynthesis and Respiration model) will be used as emission models to provide fossil fuel and biosphere data emissions

2.1.2 The EDGAR model

All the trace gases of interest share the same gridded emission model to simulate anthropogenic emissions: the EDGAR emission inventory. Originally developed by the National Institute for Public health and Environment (RIVM) and the Netherlands Organization for Applied Scientific Research (TNO), it is now in the responsibility of the Joint Research Center of the European Commission in collaboration with the Netherlands Environmental Assessment Agency (the former environmental section of RIVM).

EDGAR has a resolution of 10x10 km with global coverage and it is available on annual basis for the period 1970-2012. The database allows calculating emissions by country and sector. The sectors of interest are:

1. Energy
2. Industry combustion and process emissions
3. Aviation
4. Ground transport
5. Domestic shipping
6. International shipping
7. Residential and other combustion
8. Fuel production
9. Solvents
10. Agriculture without savannah burning
11. Solid waste
12. Wastewater handling
13. Other (fossil fuel fires, etc…)

For each tracer it is possible to obtain maps of the total emissions or specific sectors. Currently efforts are made to further improve the database to allow discrimination according to fuels. This kind of information is very important to improve the knowledge of the uncertainty's correlation between CO and CO_2 as the combustion of different fuels results in a different mix of tracers.

2.1.3 The VPRM model

Modeling of CO_2 needs to take into account the contribution of the biosphere. In the present thesis the Vegetation Photosynthesis and Respiration Model (VPRM) will address this issue. The VPRM is a diagnostic model that assimilates data from remote sensing, meteorology and flux towers to derive biospheric flux distributions at high spatial (10x10 Km) and temporal (hourly) resolution. It is also used to provide the prior knowledge for the Bayesian inversion for CO_2.

Remote sensing data comes from the MODIS (MODerate resolution Imaging Spectroradiometer) sensor carried by the NASA Terra and Aqua satellites and consists of eight-day means estimates of EVI (Enhanced Vegetation Index) and LSWI (Land Surface Water Index). EVI is used to assess ecosystem production, while LSWI is more useful in capturing effects of water stress and phenology on plant photosynthesis respectively. VPRM has a simple mathematical structure and calculate vegetation's NEE (Net Ecosystem Exchange) as a sum of GEE (Gross Ecosystem Exchange) and R (Respiration). The full model equation is given by the equation in Fig. 2.2:

$$NEE = \underbrace{\overbrace{-\gamma \times T_{scale} \times \underset{\substack{\uparrow \\ \text{LSWI}}}{P_{scale}} \times \underset{\uparrow}{W_{scale}}}^{\varepsilon} \times \frac{1}{(1+PAR/PAR_0)} \times EVI \times PAR}_{\text{GEE}} + \underbrace{\alpha \times T + \beta}_{\text{R}}$$

Figure 2.2: VPRM model mathematical structure

Where T is the temperature of the soil and ε is the light use efficiency, with γ representing the maximum quantum yield and T_{scale}, W_{scale} and P_{scale} are scalars for temperature, canopy water content and leaf phenology respectively that reduce light use efficiency. The model requires four parameters per vegetation type: α, β, γ and PAR_0, the half saturation value. A detailed description of the model is given in Mahadevan et al. (2008).

2.1.4 Introduction to the Bayesian inversion

Atmospheric inversions provide an estimate of the distribution of sources and sinks over the domain's surface from available concentration measurements ("top-down" approach). This can be formalized in the following linear relation:

$$\boldsymbol{y}=\mathbf{K}\boldsymbol{\lambda}+\boldsymbol{\varepsilon} \qquad (2.2)$$

Where the y vector contains a certain number *n* of observations, while K is the Jacobian matrix that relates the observations with the state vector λ. In the present thesis the focus will be on surface-atmosphere gas exchanges due to biospheric processes or anthropogenic emissions, so the observations are tracer's mixing ratios at different times and locations, **K** is a transport operator that maps fluxes at different times and locations and $\boldsymbol{\lambda}$ contains *m* scaling factors for the categories of interest. In the frame of this work, the **K** matrix itself is a product of two further matrices: **H** and **F**, where **H** accounts for the transport (i.e. linking gridded fluxes to mixing ratios), while **F** re-scale **H** to match the state vector (i.e. converts the state vector to gridded fluxes). The state vector accounts for seasonal fluxes, emission sectors and fuels for each one of the three trace gases from the EDGAR emission model, plus seasonal fluxes and vegetation categories for both GEE and R modeled by VPRM only for CO_2. The error vector (ε) accounts for measurement error related to uncertainty in the observation and to model-data mismatch resulting from model uncertainty.

Bayesian inversion puts measured data (MOZAIC/IAGOS profiles) and a priori information (fluxes from the emission models) together to reconstruct the most probable state vector. Gaussian distribution is assumed for the error vector, which implies random errors with no bias (Fig. 2.3).

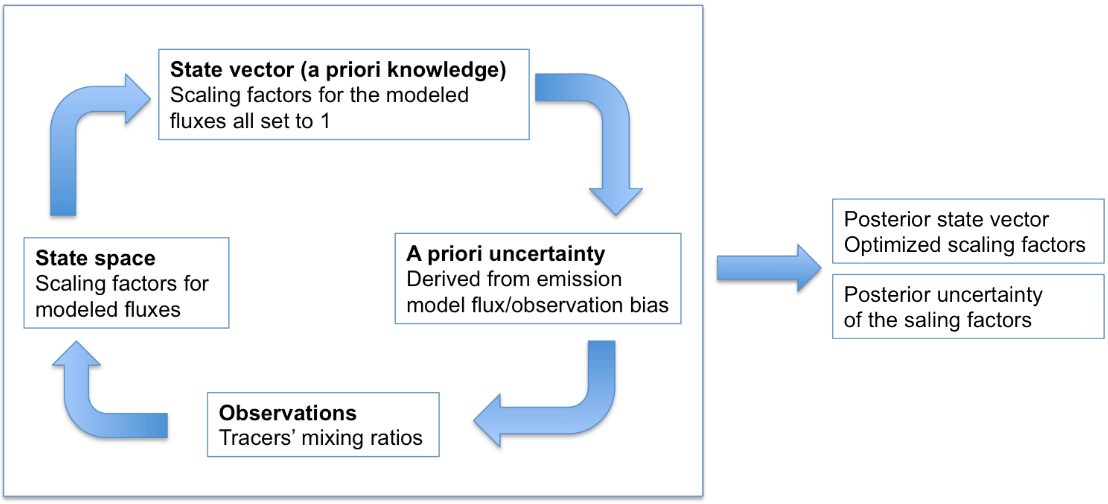

Figure 2.3: Bayesian inversion scheme

Optimum posterior estimates of the scaling factors are obtained by minimizing the following cost function J (Rodgers, 2000):

$$J(\lambda) = (y - K\lambda)^T S_{\varepsilon}^{-1}(y - K\lambda) + (\lambda - \lambda_{prior})^T S_{prior}^{-1}(\lambda - \lambda_{prior}) \qquad (2.3)$$

Where the first and the second term of the RHS are the observational constraint and the prior constraint term respectively. S_{ε} is the error covariance matrix for the anthropogenic and biogenic signals and accounts for instrumental uncertainty, transport model's error and other sources of error like boundary conditions and flux aggregation not accounted for through state vector's adjustment. Similarly, S_{prior} is the error covariance matrix for the prior scaling factor and its implementation requires a different approach for biospheric and anthropogenic fluxes. For the biosphere, it is derived from the mismatch between the modeled biospheric fluxes and the observed eddy covariance fluxes, while for anthropogenic emissions a detailed knowledge of the emission inventories is required. By construction, the error is reduced by the prior error estimate. It is also worth to note that both error covariance matrices have non-zero off-diagonal terms, as the error sources are not independent.

Minimizing the cost function results in an optimal posterior estimate of the state vector λ that is consistent with both the measurements and the prior model estimates:

$$\hat{\lambda} = (K^T S_{\varepsilon}^{-1} K + S_{prior}^{-1})^{-1} (K^T S_{\varepsilon}^{-1} y + S_{prior}^{-1} \lambda) \tag{2.4}$$

The error covariance matrix of the optimal posterior state (the posterior uncertainty) is given by:

$$\hat{S}_{\hat{\lambda}} = (K^T S_{\varepsilon}^{-1} K + S_{prior}^{-1})^{-1} \tag{2.5}$$

Note that this quantity is not dependent from both prior fluxes and measured mixing ratios, but only from their uncertainty and the transport matrix K. The uncertainty reduction can be defined as:

$$1 - \frac{\sigma_{post}}{\sigma_{prior}} \tag{2.6}$$

and ranges from 0 to 1. The numerator is the RMSE (Root Mean Square Error) of the posterior error covariance matrix, while the denominator is the RMSE of the diagonal elements of S_{prior}.

2.2 Sources of uncertainty in modeling airborne measurements: Observations, Modeling Framework, and Statistical Tools

2.2.1 Observations

In this study observations are collected from the MOZAIC/IAGOS fleet of commercial airliners; we made use of the available CO mixing ratio profiles. Measurement technique is described in Nedelec (2003), whereas the extensive MOZAIC CO database have been already used in different studies (Nedelec et al., 2005; Elguindi et al., 2010; Zbinden et al., 2013). Measurement precision from CO analyzer is ±5 ppbv CO for a 30 second response time (Nedelec et. al, 2003), with an accuracy of within 5%.

We considered only airports in the European domain with a significant number of observations (Frankfurt, London and Vienna) in the 2006-2011 time frame for all hours of the day; we did not use 2010 as observations are available for only six months and this may affect seasonality. In the profiles, continuous observations are averaged into 150 m intervals with each value referring to the mean height of the interval. Data were downloaded from IAGOS database (http://www.iagos.fr/). Flight tracks of commercial airliners usually extend up to 12 km of height, but here we limit the vertical extent to 4 km as the focus is on surface influence on atmospheric concentration.

We focus on the mixed layer (ML) as the part of the troposphere in which the contribution of local anthropogenic emissions to atmospheric mixing ratio is dominant. Conversely, above the ML, the atmosphere is mostly stably stratified, and CO mixing ratios depend mainly on long-range transport from distant emission sources. We assume that at two kilometers above the top of the ML, the influence of regional surface emissions is small, and we refer to this portion of the atmosphere as free troposphere (FT). Due to the difference of transport regime in ML and FT, CO mixing ratio is usually higher in ML; the difference between mixing ratio in ML and FT is referred to as CO enhancement and is used here as main indicator for the signal from regional surface fluxes (Fig. 2.4). Although due to the chaotic nature of turbulent transport and convection enhanced CO can also be found above the mixed layer, we focus here on the much stronger enhancement within the mixed layer.

There are many ways to calculate the depth of the mixed layer (z_i), but most of them are variations of the Bulk Richardson's number method or of the parcel method (Seibert et al., 2000). To establish the method of choice for the MOZAIC/IAGOS observations, we selected a sample of profiles for which it was possible to estimate the actual mixing height from the tracer's concentration profiles, and we compared the results from eight different methods with the tracer-based z_i treated as true value. The method that proved to be better in reproducing the tracer-based z_i was the parcel method with a 2 K excess temperature, which was therefore used to calculate the ML depth for each of the observed profiles.

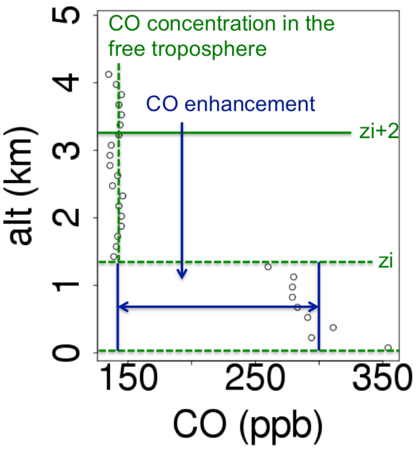

Figure 2.4: Illustration of the CO enhancement in the mixed layer (height z_i) above the CO in the free troposphere.

2.2.2: Modeling framework

The modeling framework combines a regional transport model (STILT) with an anthropogenic emission model (EDGAR) and output from a global transport model for lateral boundary conditions (MACC). For regional atmospheric transport we use the STILT (Stochastic Time-Inverted Lagrangian Transport) model, a Lagrangian particle dispersion model. Starting from each measurement (receptor) points, STILT uses analyzed wind fields from ECMWF (European Centre for Medium-range Weather Forecasts) to drive back in time for a period of ten days ensembles of simulated particles representing air parcels of equal mass (cf. Lin et al., 2003). The model uses the back-trajectories of said particles to derive sensitivity maps of the atmospheric mixing ratio measurement to the upstream surface-atmosphere fluxes (Fig. 2.5, right). By matrix-multiplication with a map of surface-atmosphere fluxes (e.g. from an emission inventory), this sensitivity map returns the simulated mixing ratio corresponding to the time and location of the observation (Fig. 2.5, left and middle). This allows for creating simulated profiles that can be analyzed in the same way as the measured ones as shown in Fig. 2.6, where a few exemplary profiles are given. In the rest of the chapter, we will refer to the Lagrangian modeling system as "STILT/EDGAR" to distinguish it from the MACC reanalysis.

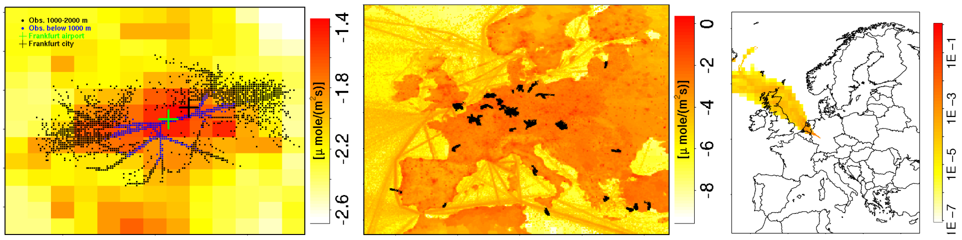

Figure 2.5: MOZAIC/IAGOS flight tracks below 4 km altitude shown on a map with CO emissions based on the EDGAR version 4.3 emissions at 10 km horizontal resolution (left), MOZAIC observation locations during 2007 in the vicinity of Frankfurt, colored by altitude (middle), and STILT/EDGAR derived footprint (sensitivity to upstream fluxes) for a single measurement location/time near Frankfurt airport (right).

Contributions to atmospheric concentrations can be either from sources and sinks close to the receptor point or from far field advection. In modeling, the former is given from the simulated fluxes in the defined horizontal domain, whereas the latter is specified by lateral boundary conditions. STILT/EDGAR derived profiles are obtained by summing the contributions from both within-domain fluxes and boundary conditions. Fig. 2.6 shows profiles from observations compared with the corresponding profiles derived from STILT/EDGAR and the boundary condition derived from the MACC reanalysis (Inness et al., 2013). The difference between the latter two should give an idea of the increase in tropospheric CO mixing ratio due to the simulated fluxes close to the measurement locations. From the figure is possible to infer that there is indeed an increase from the boundary condition in the lower part of the profile, with only exception the rightmost panel in which the STILT/EDGAR profile is indistinguishable from its boundary condition. In both modeled and observed profiles, by subtracting the free tropospheric CO value from the corresponding value in the mixed layer, the region of influence is limited to more recent emissions. A mean sensitivity map for the receptor points in the mixed layer and free troposphere, is presented in Fig. 2.7, together with the sensitivity for mixed-layer enhancements computed as the difference between the above mentioned maps for mixed layer and free troposphere.

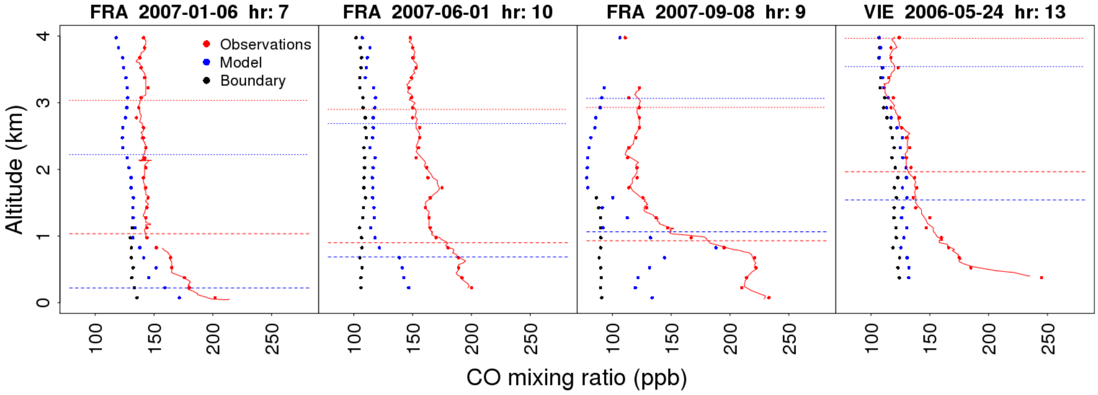

Figure 2.6: Vertical profiles from MOZAIC/IAGOS observations (red), STILT/EDGAR simulations (blue) and boundary conditions from the MACC reanalysis (black) for different locations and times. Note that observations have been plotted both as continuous data (continuous red line) and averaged over 150 m intervals (red dots). The dashed lines indicate the value of z_i and z_i+2 km for the observed (red) and modeled (blue) profile respectively.

We use STILT/EDGAR for a regional domain that covers most of Europe with a spatial resolution of 1/8 deg. latitude and 1/12 deg. longitude, corresponding to 10 km (Fig. 3.2, left). As lateral boundary condition for CO mixing ratios the MACC reanalysis (downloaded from http://www.ecmwf.int) was used.

Photochemical loss due to reaction with OH and production from CH_4 oxidation were implemented following Gerbig et al. (2003b). More detailed, the contribution to modeled CO from the advected lateral boundary condition was subject to photochemistry, but no chemical loss was included for the contribution from regional emissions, which was dominated by recent input close to the observation location with an age of typically a few hours.

For fossil fuel emissions we use EDGAR (Emission Database for Global Atmospheric Research); specifically we follow the approach taken in the COFFEE (CO_2 release and Oxygen uptake from Fossil Fuel Emission Estimate) (Steinbach et al., 2011) dataset by combining EDGARv4.3 annual global emission maps at 0.1 deg. spatial resolution for the base year 2010 provided by EDGAR (EDGARv4.2 and 4.1 are available under http://edgar.jrc.ec.europa.eu), specific for IPCC emission categories and fuel types, based on IEA (2014) fuel consumption data and EMEP/EEA (2013) emission factors. In the model we use specific temporal factors (seasonal, weekly and daily cycles) for different emission categories, and with country and fuel type specific year-to-year changes for different fuel types at national level from the BP statistical review of World Energy 2014 (BP, 2014). The basic difference to COFFEE is that here the focus is on CO rather than on CO_2 and on oxygen. This resulted in hourly resolved CO emissions, which were projected to the STILT/EDGAR EU domain. Wind fields from ECMWF have a spatial resolution of 0.25 deg. with 61 vertical levels, and a temporal resolution of 3 hours. As already mentioned in the introduction, model-data mismatch can include different aspects in both the horizontal and vertical domain, for example spatial and temporal resolution of the modeled fluxes, poorly represented convective transport in the boundary layer or biased fluxes in the emission inventory (Sect. 1.2).

Even though the main focus here is on characterizing error contributions from different sources, we also deem particularly important to investigate the effect of spatial resolution of simulated fluxes on model-data mismatch. As the STILT/EDGAR simulations are of course affected also by the other sources of error, we need to implement specific corrections accounting for the other dominant sources of error in order to single out the representation error. The following sections (2.2.3 – 2.2.7) describe the implemented corrections for the mismatch in observed and simulated depth of the mixed layer, and for accuracy of the emission fluxes.

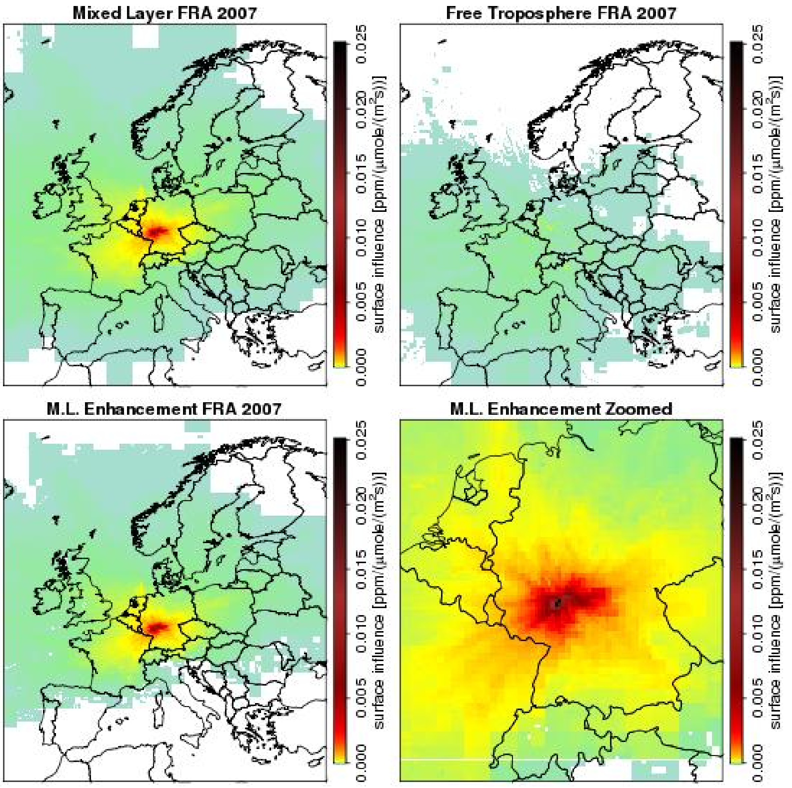

Figure 2.7: Illustration for STILT-derived mean surface influence for receptor points collected near Frankfurt in 2007 for the Mixed Layer (top left) and Free Troposphere (top right). The bottom panels show the Enhancements from the Troposphere (the difference between the former two), with different zoom.

3.1.3 Mismatch in mixing height

The depth of the mixed layer (z_i) is a very important variable in atmospheric modeling. In fact, in a one-dimensional model, the change in atmospheric mixing ratio of a trace gas due to underlying emissions is directly proportional to the ratio between emission flux and z_i (apart from the minor influence from the change in air density with altitude). Even assuming perfectly simulated fluxes, if the model returns a z_i that is higher (or lower) than the observed one, the simulated tracer in the mixed layer will be too diluted (or too concentrated), leading to a net underestimate (overestimate) in the mixing ratio enhancement. Thus tracer enhancements within the ML directly depend on the thickness of the mixed layer itself; the same emission will lead to a larger enhancement the lower the ML depth z_i is.

STILT diagnoses z_i values from ECMWF's meteorological fields using the Bulk Richardson's number method (Lin et al., 2003). The comparisons show that in general, STILT-derived values for z_i are lower than the corresponding values diagnosed from MOZAIC/IAGOS meteorological profiles. To account for this effect we apply a first order correction to the simulated enhancements that adjusts modeled z_i while maintaining the column-integrated tracer amount (Kretschmer et al., 2012). Such correction is specific for each different profile and is applied only to profiles in which the simulated z_i value exceeded 225 m (the second vertical level in the MOZAIC/IAGOS profiles). When this condition is not met in fact, the uncertainty in the observed z_i itself is expected to be too high to justify the use of a correction that adjusts the modeled value to match the observed one. Note that this situation of a simulated mixed layer height lower than 225 m occurs predominantly at night or during winter time.

2.2.4 Flux error

Emission inventories are widely recognized as important tools for atmospheric modeling. The estimated fluxes they provide are coupled with atmospheric transport in order to simulate mixing ratios that can be compared with observations. Note that spatial distributions of population density and economic activities are often used as a proxy for emissions to downscale national emission inventories. For example, an urban population gridmap, different road maps (for the 4 different types of streets), and different international aviation maps (at 3 different heights) of Janssens-Maenhout et al. (2013) are used in this paper for gridding the CO emissions in the region around Frankfurt. Such downscaling process is not perfect, and can result in biased fluxes. Differences between simulated and observed mixing ratios can then be used to quantitatively assess the emission inventories (Stein et al, 2014). In fact, as boundary condition profiles are relatively constant with height, the lion's share in the CO enhancement is accounted for by regional emissions from the emission inventories. Hence, the difference between observed and modeled enhancements reflects the difference between actual and estimated emissions.

Here we do something similar, but we take the investigation one step further. After applying the correction for mixed layer height mismatch, we assess to which degree the emission inventory correctly simulates the emission fluxes by deriving scaling factors representing the ratio between observed and modeled CO enhancements. Secondly, we apply the obtained scaling factors to correct the model output. Using these flux-corrected CO enhancements to calculate the residuals between model and observations, we remove the flux-bias contribution to the model-data mismatch, which allows for singling out the spatial representation error. To better describe this bias, observed CO enhancements are fitted against modeled enhancements using a non-linear regression model that involves three scaling factors:

$$CO_{enh,obs} = sf_{all} \cdot sf_{loc} \cdot sf_{month} \cdot CO_{enh,stilt} + \varepsilon \tag{2.7}$$

Here sf_{all} is an overall scaling factor and represents the bias component in the flux error. Conversely, sf_{loc} is specific for different airports (Frankfurt, London and Vienna) and thus represents spatial variations in the scaling factor while sf_{month} varies according to the month and allows for adjusting the seasonality of anthropogenic emissions; together, they introduce a random component in the flux error. Weighted least-squares are used to estimate the scaling factors and their uncertainties; a random error term is here indicated by ε.

Note that fossil fuel emissions are typically characterized by large areas with low emissions and small areas with strong emissions, which leads to a log-normal distribution of the enhancements. However, a least-squares optimization of the scaling factors requires a normal distribution of the dependent variable. To account for this effect, modeled and observed enhancements were log-transformed before the optimization of the scaling factors. As furthermore the log-transformation does not allow negative enhancements, the analysis was limited to the central 80% of the CO enhancements. This method can be affected by biases related to the representation of photochemistry and of the lateral boundary condition. To assess this, we performed two additional STILT/EDGAR simulations: one without taking into account any photochemistry, and one using a flat (zero) lateral boundary condition instead of MACC reanalysis fields for CO.

2.2.5 Representation error estimate

Spatial variations in emissions or fluxes at scales not resolved by a given tracer transport model are responsible for at least a large fraction of the spatial representation error (cf. Gerbig et al., 2003). Principally such representation errors can be estimated from comparisons of simulations made at different spatial resolutions. By using a Lagrangian transport model, the grid scale at which transport is combined with the emissions is flexible. STILT has a feature that allows for transport-flux coupling at resolutions of *n* times the highest resolution of 10 km (the resolution of the emission inventory); here n can assume values of 1, 2, 4, 8, 16 and 32. As coarser resolved fluxes are the result of averaging over highly resolved fluxes, the effects from localized strong emission sources are reduced with decreasing spatial resolution. The representation error can thus be written in a general way as:

$$CO_{repr.err}(n\cdot 10km) = \sigma(CO(n\cdot 10km) - CO(10km)) \tag{2.8}$$

Here *CO* is the simulated CO enhancement after correction for mixing height mismatch and flux error.

For each location and time, the difference between the simulated CO mixing ratio at the highest and a lower resolution can be interpreted as a single realization of the representation error for a specific spatial scale. Because the representation error is defined as the standard deviation of several realizations, the data need to be divided into different groups such that the representation error is not estimated as a single number, but specific for different situations. As the CO enhancements for the different airports show a strong dependence on wind direction (see section 3.4), we decided to group into 30° circular sectors. In addition, as wind direction determination is difficult at low wind speed, a further group “low wind” was formed for wind speeds of less than 3 m/s.

As the representation error is also expected to be in some way proportional to the enhancement (in the sense that larger enhancements are associated with larger errors), we estimate a relative representation error for each airport and wind sector. For this, the airport and wind sector specific data were sorted by the simulated enhancement and grouped into ten bins of equal size. For each bin, we calculate the standard deviation of the error realizations. The random component of the relative representation error is then derived as the slope of a linear model fitting the within-bin standard deviations of the realizations against the median enhancements of the bins. The bias component was computed in a similar way, but using within-bin median instead of the standard deviation. That way each profile is associated with a relative representation error that varies between airports and wind sector. The random component corresponds to the noise in the representation error whereas the bias component represents a systematic error.

The presence of a bias component for the high-resolution STILT/EDGAR simulation after applying the correction for the flux error may seem surprising, however it is to some degree expected: due to the lognormal distribution of the enhancements the bias correction is unbiased only for the log-transformed values, not for the enhancements themselves. The absolute (as opposed to relative) representation error for each profile is then derived as the product between the relative representation error and the simulated CO enhancement of the considered profile.

2.2.6 Representation error validation

In order to evaluate to which degree the estimated representation error can be useful to describe and ideally predict model-data mismatch for an independent model, the representation error was compared with residuals between the MACC reanalysis and observations from MOZAIC/IAGOS. Here we assess the dependence on wind direction and on time (month) for the random component. This was done in order to validate whether or not the representation error has any capability to describe spatial or temporal variations in model-data mismatch as a whole. The validation analysis was limited to the city of Frankfurt due to the better data coverage.

The slope of the linear regression indicates the fraction of variance and bias in the model-data mismatch that is accounted for by representation error and was derived using the Theil-Sen estimator. Such method calculates the median of all the slopes of the lines passing through a couple of points in the graph and is therefore less sensitive to outliers. Validation results are shown only for a single spatial resolution of the STILT/EDGAR simulation (80 km). This resolution was chosen, as it is closest to the MACC reanalysis horizontal resolution of 1.125 deg.

2.2.7 Contributions from different error categories

After the different sources of error (mismatch in the mixed layer depth, accuracy of the emission fluxes, and spatial representation error) contributing to model-data mismatch have been quantified, they can be compared to each other in order to derive a quantitative estimate of each source's contribution and thus to assess their relative importance. For mixed layer height mismatch and emission accuracy, the random and bias components are derived separately from mean and standard deviation respectively of the relative residuals between simulated CO enhancements before and after corrections according to equation 2.9 and 2.10 respectively.

$$R.C. = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{CO_{enh,corr} - CO_{enh,uncorr}}{CO_{enh,corr}}\right) \tag{2.9}$$

$$B.C. = \sigma\left(\frac{CO_{enh,corr} - CO_{enh,uncorr}}{CO_{enh,corr}}\right) \tag{2.10}$$

In assessing the contributions from the corrections, we consider the modeled values before the correction as stronger biased compared to the corresponding corrected values (in other words, corrected values are expected to be closer to the truth). The calculation was performed separately for both STILT/EDGAR and MACC. Contribution from representation error for each of the considered resolutions was derived from the mean of both random and bias component of representation error.

2.3 Multi-species inversion: Input Data, Modeling Framework, and Bayesian Inversion

Before describing the modeling framework, we introduce some specific terminology to reduce ambiguity in Sect. 2.3.1-2.3.6. Quantities that can be observed are termed *species*, or *trace gases*, corresponding in this case to total CO_2, CO and CH_4. These three species are simulated using five *modeled species*, namely CO_2 from fossil fuels, CO_2 related to GEE (Gross Ecosystem Exchange) and to respiration, CO, CH_4. Modeled species related to anthropogenic emissions are modeled as the sum of contributions from different *emission sectors* (Table 2.1) and *fuel types* (Table 2.2); as a further factor of discrimination, both anthropogenic and biospheric contributions are split into monthly contributions. Simulated fluxes specific for different modeled species, emission sectors, fuel types and months of the year are called *flux categories*.

	Adj IPCC	Description	Aggregated
1	1a1a	Power generation	Energy
2	1a1bcr	Other transformation non-energy use	Energy
3	1b1	Solid fuels production	Energy
4	1b2abc	Gas flaring	Energy
5	1b2ac	Oil prod., distribution and flaring	Energy
6	1b2b	Gas production and distribution	Energy
7	1a3a+1c1	International and domestic aviation	Transport
8	1a3b	Road transport	Transport
9	1a3ce	Non-road ground transport	Transport
10	1a3d+1c2	Inland waterways and shipping	Transport
11	1a2+6cd	Industrial combustion (non-power)	Industry
12	2a	Cement and lime production	Industry
13	2befg+3	Chemical industry and solvent	Industry
14	2c	Metal industry emission	Industry
15	1a4	Buildings	Buildings
16	4a	Enteric fermentation in agriculture	Agriculture
17	4b	Manure management	Agriculture
18	4c	Rice cultivation	Agriculture
19	4f	Agricultural waste burning	Agriculture
20	6a	Solid waste disposal in landfills	Waste
21	6b	Wastewater treatment	Waste
22	7a	Fossil fuel fires	FF_fuels

Table 2.1: Specific emission sectors accounted for in the state vector and aggregated categories as used in Fig. 4.3.

	Fuel type	Aggregated fuel type
1	Brown coal	Coal
2	Hard coal	Coal
3	Peat	Coal
4	Gas derivatives	Gas
5	Natural gas	Gas
6	Heavy oil	Oil
7	Light oil	Oil
8	Solid waste	Waste
9	Venting and flaring	Oil
10	Other (*)	Other
11	Gas biofuels	Bio
12	Liquid biofuels	Bio
13	Solid biofuels	Bio

Table 2.2: Specific fuel types accounted for in the state vector and aggregated categories as used in Fig. 4.4.
(*) The category "Other" is derived by summing the contribution from those processes in which is difficult to establish the specific fuel responsible for the emissions.

2.3.1 Vertical profile input data

In this study the modeled profiles have the identical structure to those collected from the IAGOS fleet of commercial airliners. The spatial and temporal coordinates of different observations will be used as input for the modeling framework whereas the observed values of atmospheric mixing ratios of CO and meteorological parameters themselves will play a role in calibrating the modeling framework.

Central for this work is the concept of the Mixed Layer (ML), the lower part of the troposphere in which trace gases are well mixed due to turbulent convection in the time scale of an hour or less, and in which the effect of regional surface-atmosphere fluxes is the strongest. As input to the inversion we use the enhancement of the species' mixing ratio within the mixed layer relative to that in the free troposphere (FT), similar to the approach described in Sect. 2.2. This mixed layer enhancement best reflects the influence of regional fluxes. To compute this, we take the average mixing ratio within the mixed layer and subtract the value taken at 2 km above the mixed layer top (z_i), i.e. well within the free troposphere. The z_i is a very important parameter in atmospheric modeling, and accounts for most of the transport uncertainty in the vertical domain. In fact, assuming that the mixed layer is the part of the troposphere in which trace gases are well mixed due to turbulent convection, given a certain amount of trace gas in the ML, its mixing ratio depends strongly on its depth z_i. Even if the model has correctly reproduced the amount of trace gas in the real mixed layer, if the modeled z_i is lower (higher) than the actual one, then the simulated ML mixing ratio will be higher (lower) than it actually should be. In the present study, modeled z_i are corrected following the method described in Sect 2.2.3.

2.3.2 Transport-flux coupling

The modeling framework is composed of a regional transport model (STILT), the EDGAR (Emission Database for Global Atmospheric Research) emission inventory to model anthropogenic emissions, VPRM (Vegetation Photosynthesis and Respiration Model) to model emissions from the biosphere and output from global transport models for lateral boundary conditions for the different modeled species. The expressions 'anthropogenic emissions' and 'fossil fuel emissions' are considered synonymous and are used to indicate the sum of fossil fuel and biofuel emissions, without including contributions from LULUCF (Land Use, Land use Change and Forestry).

For regional transport we make use of the LPDM STILT (Stochastic Time-Inverted Lagrangian Transport) (Lin, 2003) to derive the sensitivity of the atmospheric mixing ratio measurement to upstream surface-atmosphere fluxes, so-called "footprints". For each measurement location and time (also called receptor point), a single footprint is derived; this is then matrix-multiplied with an emission map from an emission inventory, resulting in a simulated mixing ratio corresponding to the regional contribution at the measurement location. A detailed description of STILT is given in Lin (2003) and Gerbig (2003). We use STILT coupled with emission models for both anthropogenic (EDGAR) and biosphere (VPRM) fluxes on a regional domain that covers most of Europe (33° to 72° N, -15° to 35° E) with a spatial resolution of 1/8 degree for latitude and 1/12 degree for longitude, roughly corresponding to 10 km. As lateral boundary condition for CO mixing ratios the MACC reanalysis (Inness, 2013, downloaded from http://www.ecmwf.int) was used, whereas for CO_2 and CH_4 we use output from the Jena CarboScope (Rödenbeck, 2003; CO_2 data available from www.bgc-jena.mpg.de/CarboScope/) which is based on forward simulations of global-inversion optimized fluxes with the TM3 transport model (Heimann and Körner, 2003). TM3 fields have lower resolution, but they are chosen for their consistency with measurements from the ground-based network. In addition, spatial resolution is of relatively minor importance for the contribution from the lateral boundary as it is far away from the measurement locations.

For fossil fuel emissions we use a model based on the EDGAR emission inventory modified following the same approach taken for COFFEE (CO_2 release and Oxygen uptake from Fossil Fuel Emission Estimate) (Steinbach et al., 2011; Vardag et al., 2015). To obtain hourly resolved emissions from the original EDGAR annual fluxes for different emission categories we add specific temporal activity factors (Denier van der Gon, 2011) to account for differences in emissions due to seasonal, weekly and daily cycles. In addition, the different emission categories are further split into contributions from different fuel types from British Petroleum's Statistical Review of World Energy 2014 (BP, 2014). The World Energy Outlook from IEA as alternative source of information was not chosen, as the report from BP is available earlier (April vs. November of the following year). This allows for taking into account changes in emissions between different years. Such an emission model provides hourly resolved fluxes for each fossil fuel flux category with a spatial resolution of roughly 10 km on our regional European domain. To take into account also the contribution from the biosphere we use the Vegetation Photosynthesis and Respiration Model (VPRM). VPRM simulates realistic patterns at small spatial (10 km x 10 km) and temporal (hourly) scales and is used here to provide the a priori fluxes for biosphere-atmosphere exchange of CO_2. This model is described in detail in Mahadevan (2008).

STILT transport is driven by meteorological fields from the ECMWF IFS (12 hour forecasts twice daily at 3-hourly temporal resolution), which have a spatial resolution of 0.25 degree with 61 vertical levels. In the rest of the chapter, we will refer to the STILT/EDGAR/VPRM/MACC/TM3 combination of transport, simulated fluxes and advected boundary conditions as merely 'STILT' for simplicity.

2.3.3 Bayesian Inversion

As explained in section 2.1.4, atmospheric inversions provide an estimate of the distribution of sources and sinks over the domain's surface from available concentration measurements making use of the so-called "top-down" approach. In this study, the state vector (Eq. (2.2)) accounts for specific emission sectors (Table 2.1) and fuel types (Table 2.2) for each one of the three modeled species from the EDGAR emission model, plus gross fluxes (gross ecosystem exchange GEE and respiration R) modeled by VPRM for 5 different vegetation classes. For both anthropogenic and biospheric fluxes the temporal resolution of the state vector is monthly. The number of state vector elements per month amounts to 69 scaling factors for the different fuel- and sector-specific anthropogenic emissions for each species, and 10 scaling factors for biosphere-atmosphere exchange (respiration and photosynthesis for each of the five vegetation classes), so in total 217 scaling factors per month, or 2604 scaling factors for the full year. To avoid large memory requirements for $\mathbf{H}$ and $\mathbf{F}$ matrices, their product is directly computed within the STILT code. The random error $\boldsymbol{\varepsilon}$ accounts for measurement error related to uncertainty in the observation and to model-data mismatch resulting from model uncertainty.

Bayesian inversion combines observations (IAGOS profiles) with a priori information (scaling factors and their a priori uncertainties) to reconstruct the most probable state vector. Optimum posterior estimates of the scaling factors are obtained by minimizing cost function J detailed in section 2.1.4 (Eq. (2.3)). The detailed error structure for model-data mismatch and prior uncertainty is described in the Sect. 2.3.4.

The targeted quantities of this study are the emissions over a specific area at a specific time scale (e.g. month); those quantities can be derived from the prior and posterior state through a spatiotemporal aggregation operator **A** that allows for the conversion of scaling factors into physically representative quantities. As a spatial aggregation scale we chose a domain encircling the 50% influence in the cumulated footprint for the receptor points in the ML for the year 2011 (Fig. 2.8). The prior and posterior uncertainty of these targeted quantities (σ_{prior} and σ_{post}) is obtained by applying the aggregation operator to the respective uncertainty covariances:

$$\boldsymbol{\sigma}_{\boldsymbol{prior}} = \sqrt{\mathbf{A}^{\mathrm{T}}\mathbf{S}_{\boldsymbol{prior}}\mathbf{A}} \quad \text{and} \quad \boldsymbol{\sigma}_{\boldsymbol{post}} = \sqrt{\mathbf{A}^{\mathrm{T}}\mathbf{S}_{\boldsymbol{post}}\mathbf{A}} \tag{2.11}$$

Different versions of the aggregation operator were created for this: emissions categories are aggregated according to different fuel types (coal, oil, gas, bio, waste, other) and according to emission sectors (energy, transport, industry, buildings, agriculture, waste, fossil fuel fires).

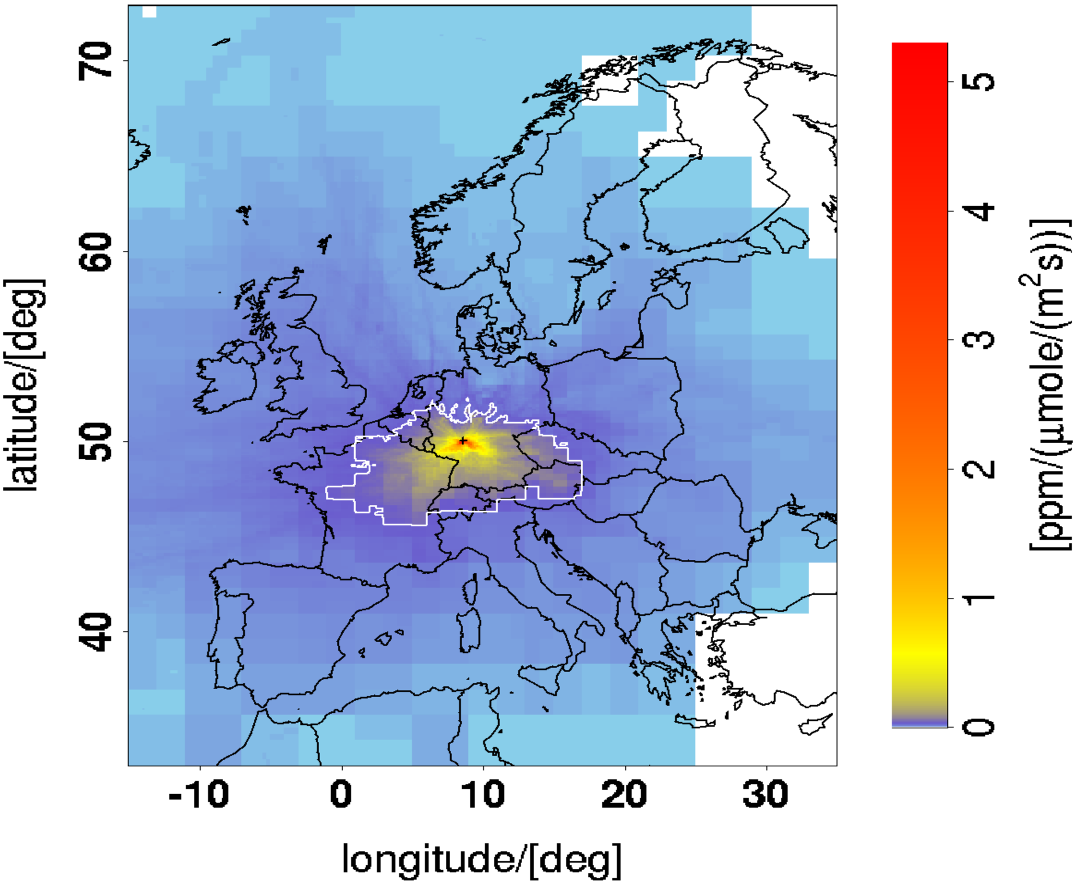

Figure 2.8: Cumulative sum of the ML footprints for all flights into or out of FRA in the year 2011. The gray line delineates the 50% footprint.

To quantitatively assess the information provided by the inversion, the reduction of uncertainty in the posterior compared to the prior estimate is a useful measure (Eq. (2.6)).

2.3.4 Prior error structure

As in this study a multi-species inversion with CO, CO_2 and CH_4 is envisioned, we have the chance to exploit the correlations in the uncertainties of the different trace gases related to both a priori fluxes and model-data mismatch. This is particularly true for CO and CO_2 because they share a larger part of the emission sources, which implies correlations in the respective uncertainties. In the multi-species inversion, such information is stored in the areas of the error covariance matrices that describe covariance between different modeled species (off-diagonal 'blocks' in Fig. 2.9b for $\mathbf{S}_{\mathbf{prior}}$ and Fig. 2.10b for $\mathbf{S}_{\boldsymbol{\varepsilon}}$). In the single-species inversions, said covariance is set to zero, corresponding to a situation where the different species are completely independent of one another. Conversely, the measurement uncertainty is stored in the main diagonal of the $\mathbf{S}_{\boldsymbol{\varepsilon}}$ (Fig. 2.10d).

We used a single year (2011) dataset restricted to the vertical profiles centered at the Frankfurt airport, and restricted to daytime during well-mixed atmospheric conditions (10:30 to 17:30 CET). The dataset contains 1098 pseudo-observations, 366 for each of the three observable species, whereas the state vector contains the scaling factors for 2604 flux categories, each equal to one in the prior.

The prior error covariance matrix can be expressed as follows:

$$\mathbf{S}_{\mathbf{prior}} = \mathbf{C}_{\mathbf{prior}} \boldsymbol{\rho}_{\mathbf{prior}} \qquad (2.12)$$

where $\mathbf{C}_{\mathbf{prior}}$ is the prior error correlation matrix (Fig. 2.9a) and $\boldsymbol{\rho}_{\mathbf{prior}}$ is a prior rescaling matrix described in Sect. 2.3.6 (Fig. 2.11a). First we describe how $\mathbf{C}_{\mathbf{prior}}$ is generated. The prior error correlation matrix is a square matrix of rank 2604, reflecting the length of the state vector, and results from the product of three components (Fig. 2.9b, 2.9c and 2.9d) accounting for correlations between flux categories according to the modeled species, emission sectors and fuel types respectively. In four different instances, a correlation of 0.7 is applied:

1. Between different anthropogenic modeled species
2. Between GEE and respiration
3. Between different emission sectors
4. Between different fuel types

Such a correlation implies that the explained variance for each constraint everything else being equal is roughly 50%, (0.7 to the square equals 0.49) with the rest remaining independent. In addition, the correlation between fossil-fuel-related and biosphere-related scaling factors is zero, and the same holds for fluxes of different months, indicating complete independence from one another.

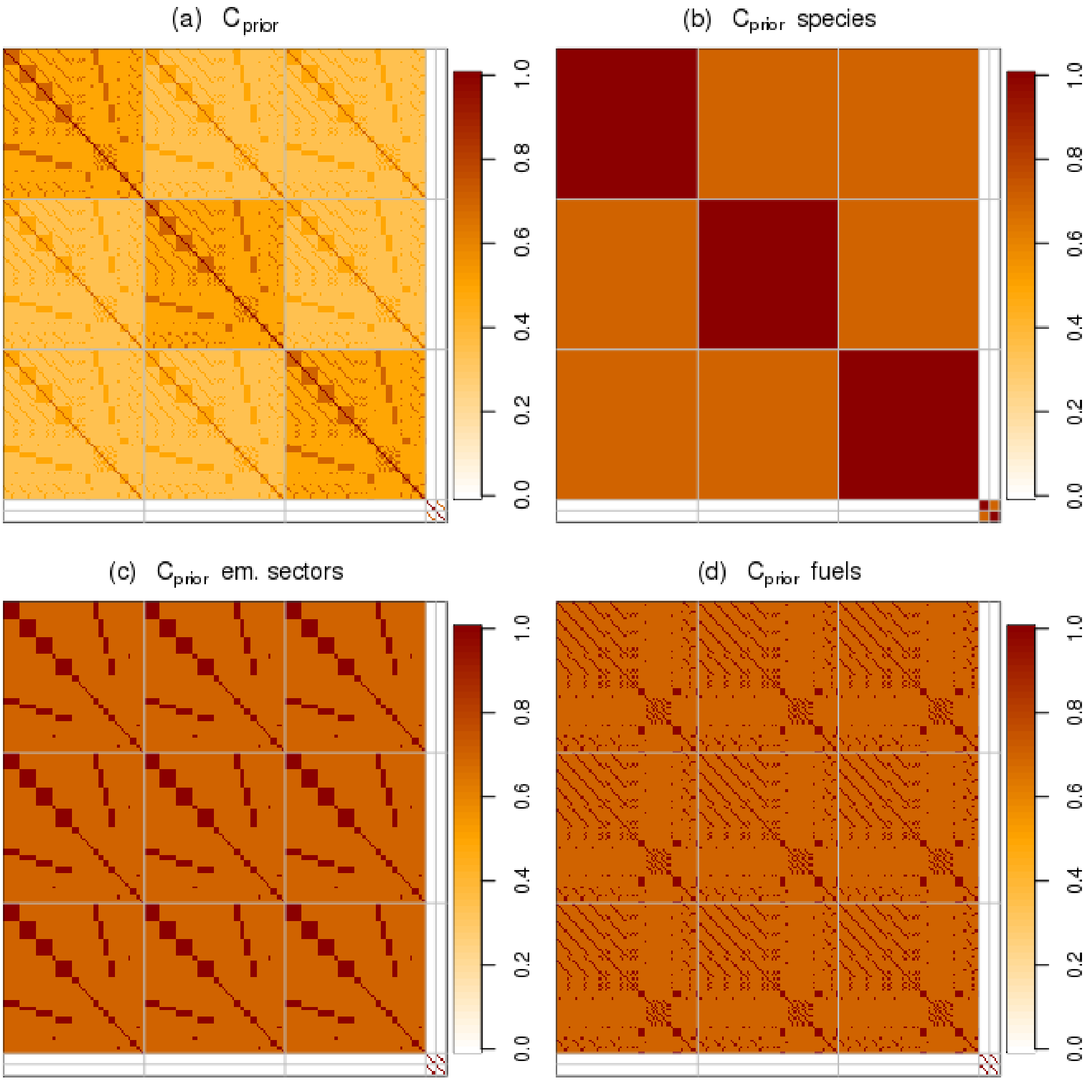

Figure 2.9: Prior error correlation matrix (a) used in the multi-species inversion, and the respective components for modeled species (b), emission sectors (c) and fuel types (d). Matrix (a) is the element-wise product of matrices (b), (c) and (d). Each matrix has the same dimensions (2604x2604) reflecting the length of the state vector. The matrices are shown for only one month here, for illustration. The gray lines indicates subsets of the flux categories according to different modeled species ('blocks'), ordered as follows from top to bottom and from left to right: anthropogenic CO_2, CO, CH_4, GEE and respiration. In the single-species inversion, the correlation values in the off-diagonal 'blocks' of matrix (b) are set to zero. In the complete matrix, correlation between fluxes from different months is also set to zero.

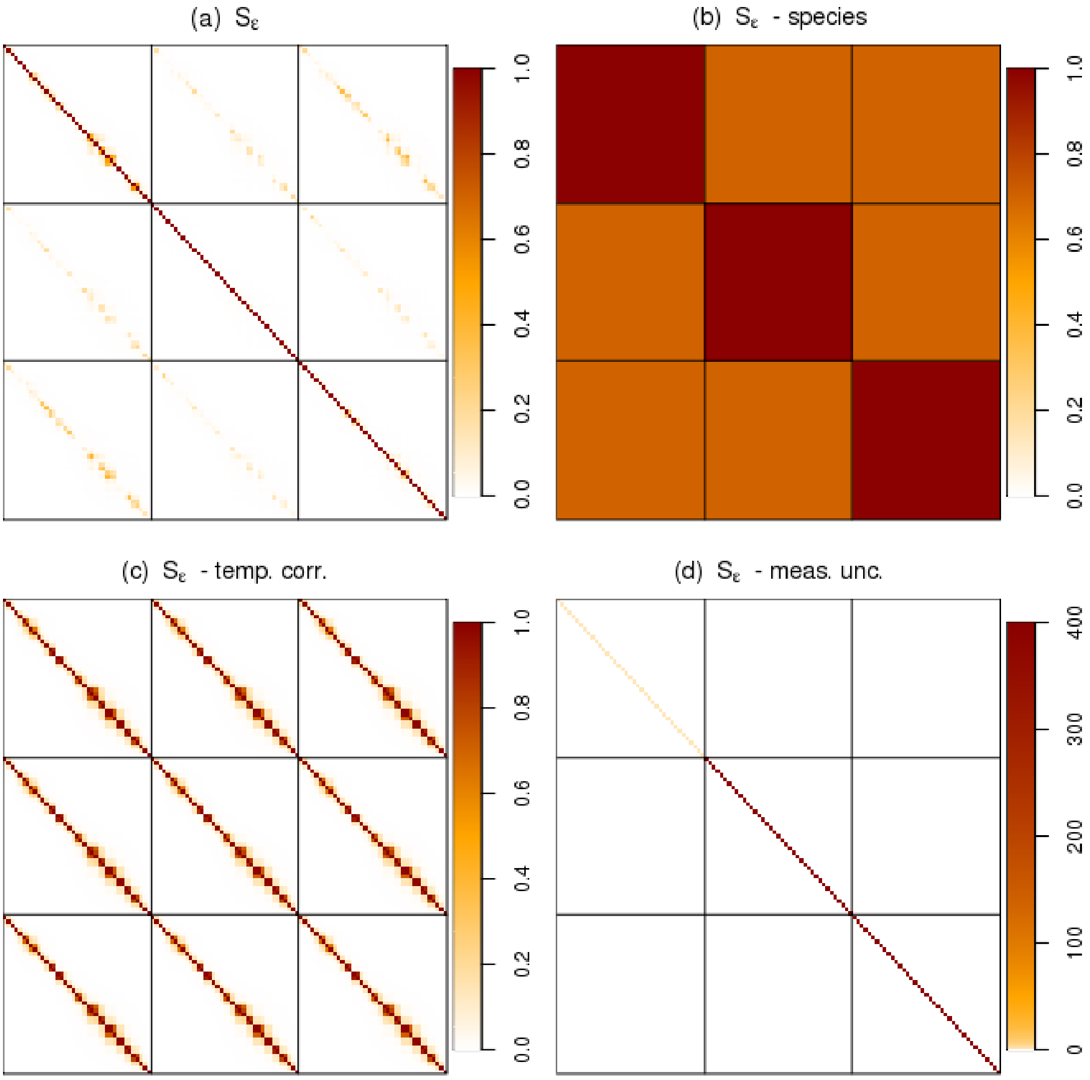

Figure 2.10: Model-data mismatch correlation matrix (a) used in the multi-species inversion, species correlation matrix S_s (b), temporal correlation matrix S_t (c) and squared measurement uncertainty (d). Note that the measurement uncertainty is expressed in ppm for CO_2 and ppb for CO and CH_4. Each matrix has the same dimensions (1098x1098) reflecting the length of the observation vector, but here only the data of July are plotted to increase visibility. The gray lines indicate different species in the observation vector ('blocks'), ordered as follows from top to bottom and from left to right: total CO_2, CO and CH_4. In the single-species inversion, the correlation value in the off-diagonal 'blocks' of matrix (b) is set to zero. The structure in S_s in (c) is a result of the uneven temporal distribution of the observations within the month.

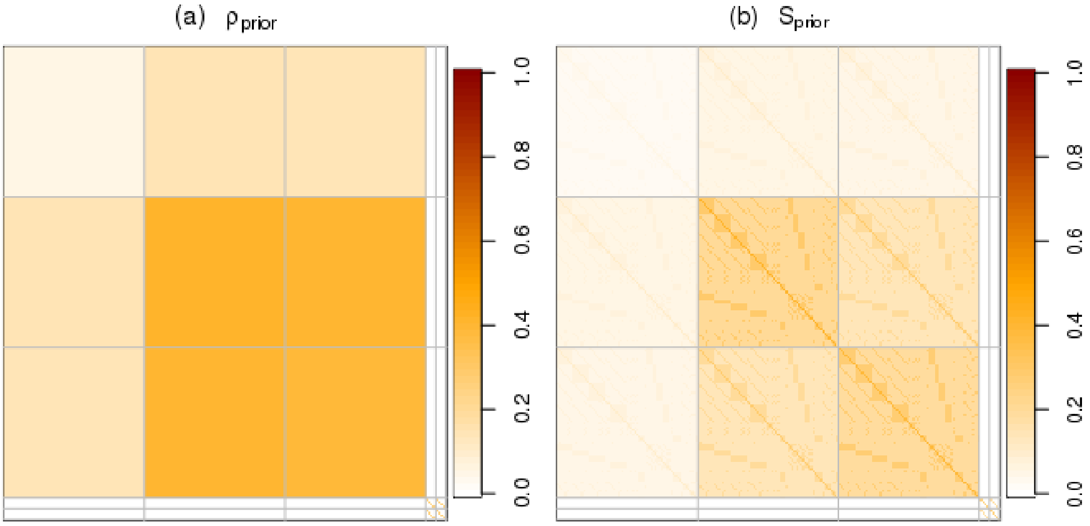

Figure 2.11: The final rescaling matrix ρ_{prior} (a) and the prior error covariance matrix S_{prior}(b). Note that ρ_{prior} can be defined as the element-wise ratio of S_{prior} and C_{prior}.

2.3.5 Prior error scaling

After having specified the prior error correlation matrix $\mathbf{C}_{\mathbf{prior}}$, we now describe how we rescale it to obtain $\mathbf{S}_{\mathbf{prior}}$; for this task we rewrite Eq. (2.12) as

$$\mathbf{S}_{\mathbf{prior}} = \mathbf{C}_{\mathbf{prior}}\boldsymbol{\rho}_{\mathbf{prior}} = \tag{2.13}$$

$$= \begin{pmatrix} \boldsymbol{C_{11}} & \boldsymbol{C_{12}} & \boldsymbol{C_{13}} & \mathbf{0} \\ \boldsymbol{C_{21}} & \boldsymbol{C_{22}} & \boldsymbol{C_{23}} & \mathbf{0} \\ \boldsymbol{C_{31}} & \boldsymbol{C_{32}} & \boldsymbol{C_{33}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{C_{bio}} \end{pmatrix} \begin{pmatrix} 1/\rho_1\rho_1 & 1/\rho_1\rho_2 & 1/\rho_1\rho_3 & 0 \\ 1/\rho_2\rho_1 & 1/\rho_2\rho_2 & 1/\rho_2\rho_3 & 0 \\ 1/\rho_3\rho_1 & 1/\rho_3\rho_2 & 1/\rho_3\rho_3 & 0 \\ 0 & 0 & 0 & \rho_{bio}^2 \end{pmatrix}$$

where each $\mathbf{C}_{\mathbf{ij}}$ is a subset of the fossil fuel part of $\mathbf{C}_{\mathbf{prior}}$ ('block') as shown in Fig. 2, and each ρ_i is defined as

$$\rho_i = \sqrt{\frac{\boldsymbol{A_i'^T}\mathbf{C}_{\mathbf{ii}}\boldsymbol{A_i'}}{\left(\sum_j A'_{ij}\,\varepsilon_i\right)^2}} \tag{2.14}$$

where **A'** is the aggregation operator for annual fluxes over the full domain, and ε_i is the corresponding relative prior uncertainty, assuming the values specified in Table 3 for different cases. Case 1 is considered as the default case, with prior uncertainty values conservatively assumed to be 20%, for CO_2 and 50% for CO and CH_4. Conversely, $\mathbf{C}_{\mathbf{bio}}$ covers the biosphere part of $\mathbf{C}_{\mathbf{prior}}$, and for $\sum A_i'\varepsilon_i$ for ρ_{bio} we use a prior uncertainty of 0.54 GtC y^{-1}, as derived in Panagiotis (2016) for the VPRM model. The biospheric part of the prior error covariance matrix assumes no correlation with the fossil fuel species.

The posterior of each Bayesian inversion depends on its specific prior. As the multi- and single-species inversion have different prior uncertainty structures, the uncertainty reduction for targeted quantities cannot be directly compared (Eq. (2.5)). To be able to compare the two inversions, we require that the a priori aggregated uncertainty of the targeted quantities remains the same, and distribute it differently each time; the prior rescaling matrix $\boldsymbol{\rho}_{\mathbf{prior}}$ is needed for this task. The benefits were tested for observations taken in different months and for three different error structures in the prior uncertainty. As a priori aggregated uncertainty we use a percentage of the aggregated modeled emissions for fossil fuels across the whole year. Table 2.3 shows the percentage values used for different cases.

	CO_2	CO	CH_4
Case 1	20%	50%	50%
Case 2	10%	50%	50%
Case 3	10%	25%	25%

Table 2.3: relative uncertainty of the prior fluxes aggregated domain-wide and annual for the different cases

2.3.6 Model-data mismatch error structure

In an atmospheric inversion, the model-data mismatch from every uncertainty source (such as measurement uncertainty, transport model uncertainty, spatial representation error due to limited model resolution, and boundary condition inaccuracies) needs to be taken into account. In our inversion scheme, we parameterize both the transport model uncertainty and the measurement uncertainty, with the latter playing a minor role. The model-data mismatch covariance matrix ($\mathbf{S}_{\boldsymbol{\varepsilon}}$) is constructed according to the following equation:

$$\mathbf{S}_{\boldsymbol{\varepsilon}} = \mathbf{C}_{s}\mathbf{C}_{\mathrm{t}}\boldsymbol{\varepsilon}_{tran}^{2} + \boldsymbol{\varepsilon}_{meas}^{2} \qquad (2.15)$$

where $\mathbf{C}_{\mathbf{s}}$ accounts for correlations between different observed species (Fig. 2.10b), $\mathbf{C}_{\mathbf{t}}$ accounts for the temporal correlation (Fig. 2.10c), $\boldsymbol{\varepsilon}_{tran}$ is the total transport error and $\boldsymbol{\varepsilon}_{\mathbf{meas}}^{2}$ accounts for all of the non-transport related errors like spatial representation error and lateral boundary conditions (Fig. 2.10d).

The assumed measurement uncertainty is 1 ppm for CO_2, 20 ppb for CO and 20 ppb for CH_4, while ε_{tran} is time dependent and assumed to be proportional to the modeled enhancement due to regional fluxes. The assumed measurement uncertainty is higher than the expected instrument precision because it also includes in addition the uncertainties related to spatial representation and lateral boundary condition. ε_{tran} is characterized as follows by different components in the vertical and horizontal domain:

$$\varepsilon_{tran} = enh\sqrt{\left(\varepsilon_{tran\_h}^2 + \varepsilon_{tran\_v}^2\right)} \tag{2.16}$$

where both the horizontal transport error ε_{tran_h} and the vertical transport error ε_{tran_v} are characterized as percentage error. ε_{tran_h} is assumed to be 50%, while ε_{tran_v} is a profile-specific relative error related to the modeled z_i. The vertical transport error accounts for the difference in vertical resolution between the transport model and the IAGOS profiles, and its mean value is about 10%. For the horizontal component, an uncertainty of 50% is a conservative estimate based on Lin and Gerbig (2005), where the horizontal transport error is found to be 5.9 ppm for CO_2. This, combined with about 10 ppm of drawdown in the mixed layer relative to the free troposphere, gives something like 50% error in the regional flux signal. The vertical component is so much smaller in percentage since the simulated mixing ratios are already corrected for mismatch between modeled and observed z_i.

In the multi-species inversion, the transport error correlation across species is 0.7 (Fig. 2.10b), while in the single-species inversion this is set to zero. Time correlation is assumed to decay exponentially with an exponential constant of 12 hours. The between-species correlation for model-data mismatch related to transport uncertainty reflects the fact that species are partially co-emitted and share the same atmospheric transport (and its related uncertainty).

2.3.7 Pseudo-data generation for the synthetic experiment

As explained in the introduction, in situ measurements are not available for all of the three trace gases of interest, but only for CO. For this reason this paper aims to evaluate the benefits of a multi-species inversion over a corresponding single-species one by performing a synthetic experiment, using pseudo-observations derived by perturbation of the model outputs based on a priori state vector values. The pseudo-observation vector is obtained by matrix multiplication between the Jacobian matrix **K** and what we assume to be the true state vector. The true state vector itself is obtained by using the sum of the prior state vector (all values equal to one) and a random realization of the prior error, truncated to avoid negative state vector values. This ensures that the difference between the true and prior state vector has the same error correlation structure as described by the prior error covariance matrix.

2.4 Lidar validation: Observations, Modeling Framework, and Statistical Tools

2.4.1 Profiles of in-situ observations

Figure 2.12 (left) shows an emission map over Europe, where the black points are locations in which the HALO aircraft took a measurement during the CHARM-F test campaign. The CHARM-F evaluation campaign took place between April 4th and May 13th, 2015; a total of 17 profiles were collected during five days of flight. Flight tracks of commercial airliners usually extend up to 11 km of height, for this reason observed profiles collected by such platform extends to the same altitude. True to its name, HALO is able to reach 15 km of height, allowing to better investigate the tropopause region and the lower stratosphere. Horizontal stretch of profiles during the CHARM-F campaign is between 100 and 400 km.

HALO in-situ observations are collected from the JIG sensor, extendedly described in Filges et al. (2015) as IAGOS-core GHG package. JIG allows for the simultaneous measurement of mixing ratio observations of CO_2, CO and CH_4, but in this paper we focused on CO_2 and CH_4, that allow for a comparison with the CO_2 and CH_4 retrieval of the CHARM-F lidar. JIG makes use of the Cavity Ring-Down Spectroscopy (CRDS) technique, described a.o. by Chen et al. (2010). Time resolution is 2.3 seconds for CO_2 and CH_4, and 3 minutes for CO, while the overall uncertainty is 0.13 ppm for CO_2, 0.13 ppb for CH_4 and 4 ppb for CO (cfr. with Table 3, Filges et al., 2015).

In the present paper, each profile is divided into three parts across the vertical domain: Mixed Layer (ML), Free Troposphere (FT), and Stratosphere (ST). Starting from the ground, ML is defined as the part of the troposphere that is well mixed by turbulent mixing in the timescale of one hour or less, and as such in which the contribution of local anthropogenic emissions to atmospheric mixing ratio is dominant. ML's depth can vary strongly accordingly to local orography, meteorological conditions and latitude, but at temperate latitudes rarely exceeds 3 km. Above the ML the atmosphere is mostly stable and stratified due to the negative vertical gradient of temperature. In addition, the atmospheric mixing ratios depend mainly on long-range transport from distant emission sources. Due to the difference of transport regime in the ML and above, mixing ratio is usually higher in the ML. This however also depends on whether there is presence of sources or sinks at the surface. The end of the troposphere and beginning of the stratosphere is indicated by the tropopause, an inversion in the vertical gradient of temperature due to absorption of solar radiation from ozone in the ST. Such inversion occurs at different heights depending on season and latitude, generally ranging between 17 km to the equator and 9 km to the poles (Gettelman et al., 2003). In the CHARM-F campaign profiles, the lowest observed altitude at which the vertical temperature gradient is inverted is 9 km. For these reasons, we hereafter refer to ML as the part of the profile between the surface and 3 km of height, ST as the part of the profile above 9 km of height, while the part of the atmosphere between ML and ST (3-9 km) is herein called Free Troposphere (FT).

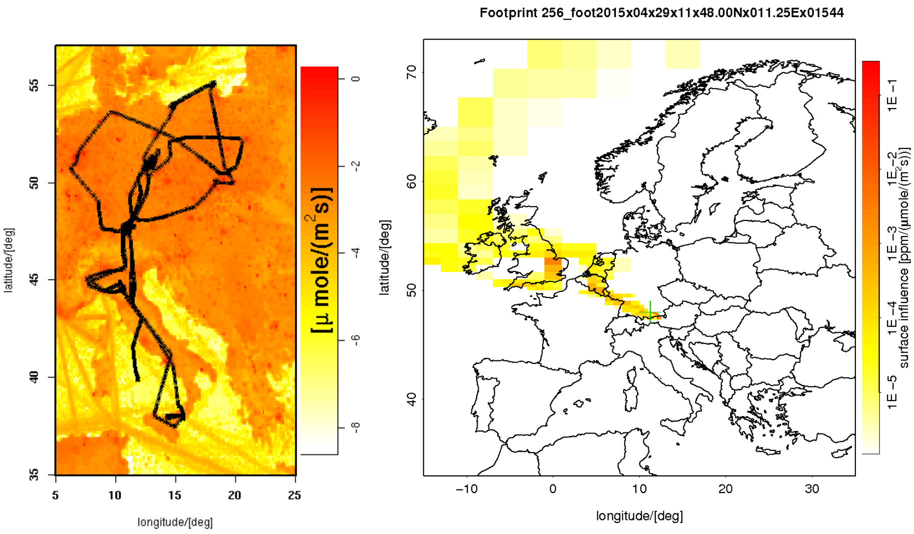

Figure 2.12: Emission map from EDGAR with receptor points of the CHARM-F validation campaign (left) and example of STILT-derived footprint for a specific receptor point (right).

2.4.2 Vertical lidar-retrieved partial columns

CHARM-F is a newly developed IPDA lidar whose distinctive feature is to achieve simultaneous measurements of CO_2 and CH_4 partial column-weighted dry-air mixing ratio below the flight track of an aircraft. This is achieved by using four lasers in pulse mode with nominal sampling rate of 50 Hz. The standard deviation of a single-shot mixing ratio is about 0.7% for CO_2 and 1.2% for CH_4. Averages up to 500 single-shot measurements, corresponding to 10 seconds and 2 km of flight, follow the inverse of square root of time, and allows for a precision below 1%. Higher precision, below 0.5% can be achieved for 100 seconds and 20 km average. Precision of the lidar ranging is about 0.25 m, while its accuracy is lower than 5 m. The IPDA technique in general and the CHARM-F sensor in particular are extensively described in Amediek et al (2017, in preparation).

2.4.3 Transport-flux coupling

The modelling framework is composed by the Lagrangian particle dispersion model STILT (Stochastic Time Inverted Lagrangian Transport), the EDGAR (Emission Database for Global Atmospheric Research) emission inventory for anthropogenic emissions and the VPRM (Vegetation Photosynthesis and Respiration Model) emission model for biogenic emissions. As STILT is a regional transport model, we derive background values from global circulation models, namely from the MACC reanalysis for CO_2, and from the TM3 model for CH_4.

The main purpose of STILT is to derive footprints, defined as the sensitivity maps of the atmospheric mixing ratio to the upstream surface-atmosphere fluxes. In each of the observation locations, also called receptor points, the model releases an ensemble of virtual particles, simulating air parcels of equal mass, driven by simulated wind fields from ECMWF. The footprint is the cumulated sum of the surface influence of the individual particles' trajectories and is to be matrix-multiplicated with specific emission maps from the emission inventories to derive the different contributions to the modelled mixing ratio in the corresponding receptor point. For a more detailed description of STILT see Lin et al. (2003) and Gerbig et al. (2003). We use STILT on a domain covering most of Europe, more precisely 33° - 72° N of latitudes and 15° W – 35° E of latitudes, and couple the STILT-derived footprints with simulated fluxes from emission inventories.

To simulate anthropogenic and biogenic emissions we use EDGAR (version 4.3) and VPRM respectively, both with a spatial resolution of 1/8 deg for longitude, and 1/12 deg for latitude, which roughly correspond to a 10 km x 10 km resolution. EDGAR simulates anthropogenic emission fluxes with annual temporal resolution, however we add specific temporal activity factors to describe variations due to the seasonal, weekly and daily cycle (Denier van der Gon et al., 2011) in order to obtain hourly resolved fluxes. In addition, compared with older versions of EDGAR, in version 4.3 also output is provided for emissions from different emission categories, split into contributions from different fuel types. The described approach is derived by similar modelling methods used for the COFFEE (CO_2 release and Oxygen uptake from Fossil Fuel Emission Estimate) dataset (Steinbach et al., 2011, Vardag et al., 2015), and provides hourly fluxes with 10 km x 10 km resolution over the previously specified European domain. To allow for converting emissions for the EDGARv4.3 reference year 2010 to other years, we use the consumption statistics at national and fuel-type level from the BP statistical review of World Energy 2014 (BP, 2014). VPRM makes use of both satellite remote sensing and eddy covariance derived parameters to model Net Ecosystem Exchange as a sum of Gross Primary Production and Respiration for eight different vegetation classes. Here we use simulated hourly biogenic fluxes at 10 km x10 km resolution. A more detailed description is available in Mahadevan et al. (2008).

As lateral boundary condition for CO_2 mixing ratio the MACC reanalysis was used (Inness et al., 2013, downloaded from http://www.ecmwf.int), whereas for CH_4 we used output from the Jena CarboScope (Rödenbeck et al., 2003, data available from www.bgc-jena.mpg.de/CarboScope/) which is based on forward simulations of global-inversion optimized fluxes with the TM3 transport model (Heimann and Körner, 2003). In the frame of this study, the STILT transport model was driven by meteorological wind fields from the ECMWF IFS (3-12 hour forecasts twice daily at 3-hourly temporal resolution), with a spatial resolution of 0.25 deg. and 61 vertical levels. To simplify the notation, in the following discussion we will refer to the STILT/EDGAR/VPRM/MACC/TM3 combination of transport, simulated fluxes and boundary conditions as merely 'STILT'.

2.4.4 Bridging JIG and CHARM-F using STILT

HALO carries many instruments, but here we are interested in two specific sensors: the JIG instrument for simultaneous in-situ measurements of CO, CO_2 and CH_4, and the CHARM-F lidar capable of parallel CO_2 and CH_4 observations. Our intent is to use the modelling framework previously described as a bridge between the two measurement systems. As only a limited amount of observed retrievals from the lidar have been made available, while still being affected by several intrumental artefacts, the main idea of this work is to calibrate the model to match the in-situ measurements, and then use the calibrated model to establish the traceability of the indirect remote sensing measurements to global calibration scales for CO_2 and CH_4 with a known uncertainty. To simulate the in-situ observation from JIG, the modelling framework is run for receptor points placed directly along the flight track, while the simulation of the lidar retrievals requires two steps. For each receptor point along the HALO's trajectory, the model first derives a vertical profile placed directly below the flight track, then the column is derived by calculating the pressure-weighted mean of the modelled mixing ratio values along this vertical profile. In this study the modelling framework is run with respect to the sampling locations showed in Figure 2.12 (left). However, not all of the receptor points where considered, but only those during ascent and descent (i.e. during the profiles).

2.4.5 Bias correction

The difference between model and observations, i.e. the model-data mismatch (MDM) is generally characterized by both systematic (bias) and random (variance) errors. In the present study we deal with MDM by following a two-steps approach: calibration of the modeling framework to better match the in-situ measurements, followed by the validation of the lidar retrievals. In calibrating the model to match the JIG in-situ observations, we specifically aim at minimizing the model-data mismatch, while during validation the main scope is instead to find a proxy with the same variance of the MDM to estimate the uncertainty of lidar retrievals.

To account for the bias between observed and modelled profiles we derive a daily correction profile to be added to the modelled profiles. Said correction profile is derived by applying a triple linear fit to the mean vertical profile of the MDM for each of the five days with flight data. The fits are specific for the MDM values in the Mixed Layer, Free Troposphere and Stratosphere. The linear fit throughout the FT is the first to be derived, and its lower and upper limits are taken as passage points for the ML and ST fits. The total number of parameters is 20, counting 3 slopes and 1 offset for each of the five day of the measurement campaign, which avoids the danger of over-fitting. Figure 2.13 shows the CO_2 correction profile for the average profile during the second day of the CHARM-F validation campaign.

It was found that the MDM in the ML is dominated by a few profiles, more precisely two CH_4 profiles in the area of Parma (Italy) and a total of four CO_2 profiles in the areas of Parma and Krakow (Poland). To investigate the impact of these profiles, both the bias correction and the error propagation analysis (cfr Sect. 2.4.6) were performed for all of the profiles and for a subset in which these high-MDM profiles were not considered.

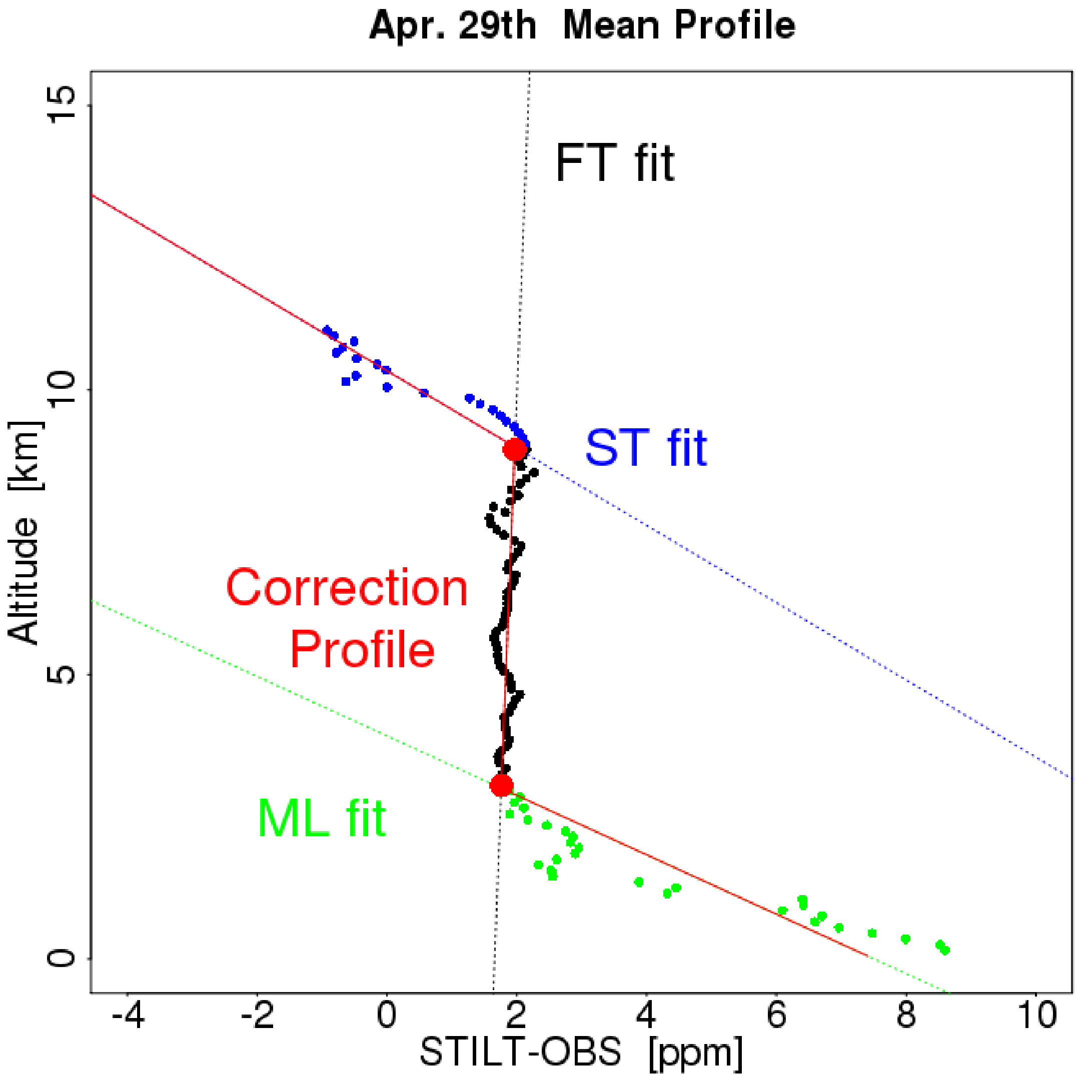

Figure 2.13: Mean residual profile for CO_2 on the second day of the CHARM-F validation campaign. The three linear fits and residuals in the Mixed Layer (green), Free Troposphere (black) and Stratosphere (blue) are also plotted. The interpolated correction profile used for the correction is plotted in red.

2.4.6 Error propagation

An estimate for the uncertainty of the lidar-retrieved vertical partial column (VPC) can be derived by propagating the model-data mismatch from the in-situ derived slanted partial columns (SPC). For each receptor point the corresponding SPC is the pressure-averaged column integrated from ground to "current" altitude using in-situ measurement along the flight track. This allows for creating profiles showing the values of SPC at different heights. Note that before calculating SPCs, the modeled profiles are corrected for bias using the calibration described in Sect. 2.4.5.

The discrepancies between observations and model output are typically the result of many different sources of uncertainty. Among these, the most important are uncertainty in simulated transport, biases in the simulated fluxes, and spatial representation error (Boschetti et al., 2015). Uncertainty in transport simulations arise from wind error (horizontal transport error) (Lin and Gerbig, 2005), but can also result from mismatch in the depth of the ML that typically arises when the model fails in simulating the height of the mixed layer (Gerbig et al., 2008). Building on the definition given in Boschetti et al. (2015), the spatial representation error (SRE) can be defined as the uncertainty arising from aggregating the emission fluxes from punctual to the spatial resolution of the emission inventory.

The representation error of each receptor point is here assumed to be the standard deviation of five mixing ratio simulations. The first of these five realizations is the output of the modeling framework run at the receptor point along the flight track, while the other 4 are the corresponding model output of the modeling framework run in receptor points artificially translated by an amount equal to a single pixel of the emission grid, namely 1/12 deg of longitude and 1/8 deg of latitude. It is clear from the previous description that to derive SRE, the whole modeling framework is involved, which implies that MDM at different heights can be approximated by SRE. By contrast, correcting only the accuracy of simulated fluxes would have an effect mainly in the mixed layer, while correcting only the boundary conditions would also affect the free troposphere.

After verifying that the dependence of both SRE and MDM on heights is similar by plotting values of SRE and MDM for slanted partial columns at different height, both SRE and MDM values were fit together in a liner model. The derived coefficients of the fitting (a scaling factor and an offset) were then used to train the representation error in order to match the model-data mismatch. For this reason the trained SRE can be used as a proxy of the model-data mismatch as a whole, and is hereafter referred to as ETU (Expected Total Uncertainty). Following these considerations, the values of the ratio between model-data mismatch and ETU are expected to be normally distributed, with mean equal zero and standard deviation equal one.

As the ETU is a proxy for the model-data mismatch as a whole, we derive an estimate of the uncertainty of VPC by calculating the representation error of the VPC and then adjust it using the same scaling factor and offset derived for the SPC. Representation error of the vertical partial column is derived in a similar fashion as the representation error of the slanted partial columns, the only difference being that the used receptor points are not corresponding to the in-situ measurement locations, but are organized in vertical arrays placed directly below the flight track.

Chapter 3 : Contribution of different error sources in the modeling of MOZAIC/IAGOS vertical profiles

As explained in the introduction, this thesis is focused on airborne in-situ measurements, which provide an alternative to more traditional atmospheric observation techniques like tall towers and low orbit satellites. The reason that makes airborne observations interesting is that collecting observations at different heights along the vertical path of an airplane allows to better describe the internal structure of the atmosphere rather than, for example, total columns from remote sensing. In addition, despite the fact that aircraft profiles do not extend to the end of the atmosphere, they still reach higher altitudes than towers, while often providing better vertical resolution than satellite profiles. For long time limited due to the cost of a rental aircraft, in the last two decades this stream of data has been gaining momentum as programs like CONTRAIL and MOZAIC/IAGOS solved the problem of costs by using a fleet of commercial airliners to provide observations on a regular basis.

This chapter is focused on describing the contributions to the model-data mismatch between observed CO mixing ratio from MOZAIC/IAGOS and the corresponding modeled values from a modeling framework. This investigation is also used to infer the representativeness of the MOZAIC/IAGOS dataset at continental scale in the European domain (cfr. Section 1.2).

3.1 Observed mixing ratios

Regular observations from MOZAIC/IAGOS allows for a thorough description of the internal structure of the troposphere. Figure 3.1 shows mean monthly values for CO mixing ratios collected around Frankfurt at four different heights each 1000 meters. A strong seasonal cycle is present in all the investigated years, with higher mixing ratios in winter-spring and lower mixing ratios in summer-fall. CO mixing ratios around Frankfurt range between 115 ppb and 205 ppb at 1000 m height and between 90 ppb and 155 ppb at 4000 m. In addition, atmospheric concentration values decrease with increasing heights. It is worth to point out that the decrease in mixing ratio between 1000 m and 2000 m is much larger than the same decrease in the 2000-3000 m and 3000-4000 m step. London and Vienna have similar patterns (not shown) but different concentration ranges. In London, CO mixing ratios range from 100 to 210 ppb at 1000 m height, and from 85 to 160 ppb at 4000 m, whereas for Vienna the range is 130-220 ppb at 1000 m and 105-170 ppb at 4000 m of height.

Abnormally high concentration values in the spring of years 2007 and 2008 at 3000 and 4000 m altitude are due to a higher occurrence of spring wildfires in many European countries compared to other years. In 2007 the affected countries were Portugal, Spain, Austria, Hungary, Germany and Czech Republic, whereas in 2008 they were Portugal, Spain, Turkey and Cyprus (Camia et. Al, 2009).

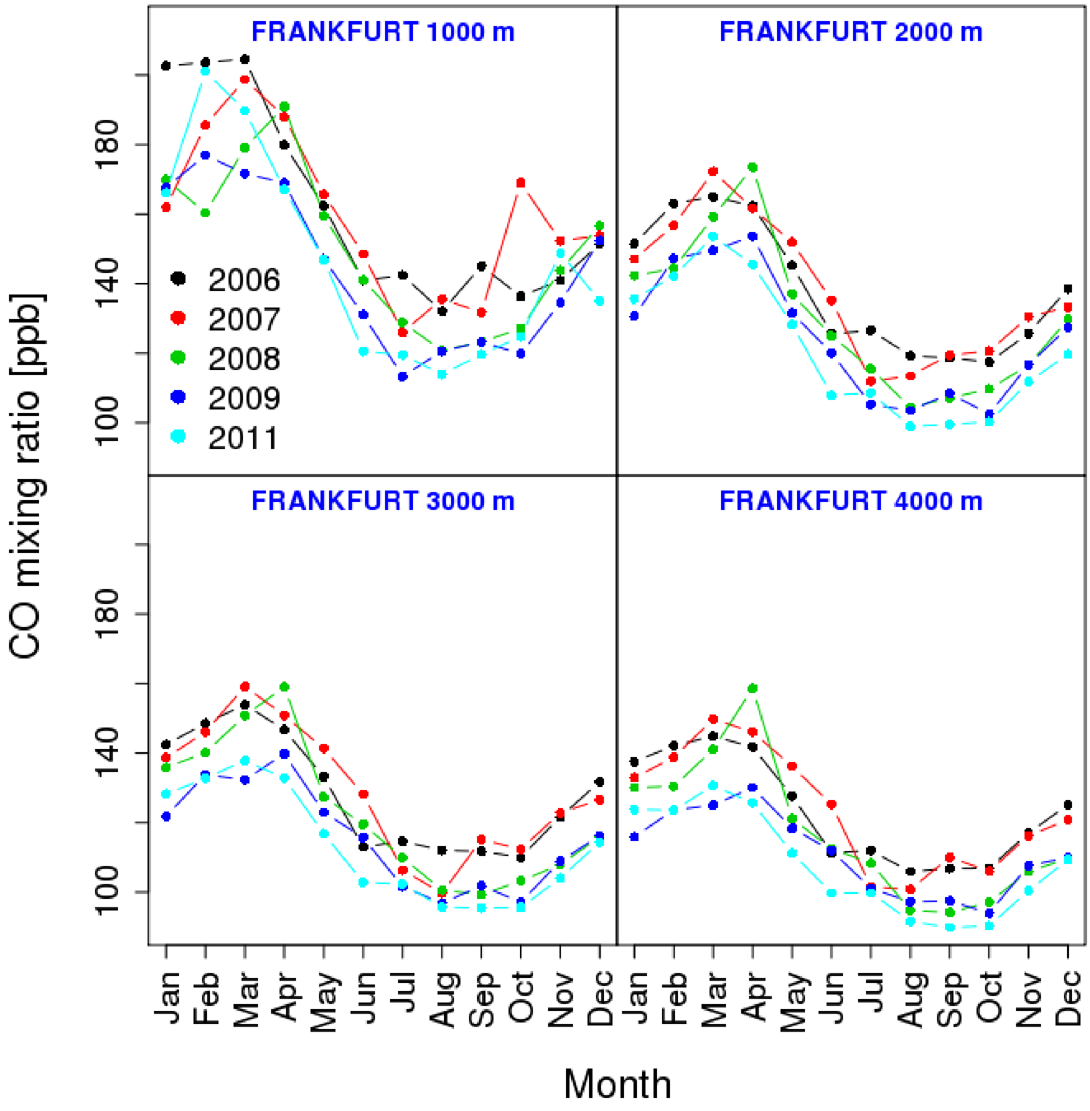

Figure 3.1: Observed CO mixing ratio for the years 2006-2011 in the lower Troposphere around Frankfurt. The plots show mean monthly values at four different heights. Note that values collected at 1000 m are strongly different than values collected at higher levels. In the spring (March and April) of 2007 and 2008, higher values were collected at 2000 m and above. This is likely due to an unusually high number of spring wildfires in many European countries.

After presenting some observational results, a comparison between observation and model time series may be helpful to introduce some general remarks. Fig. 3.2 shows an overview over four years (2006-2009) of profiles between the surface and 4 km over Frankfurt. It is clear by comparing the top panel with the two middle panels that both models correctly simulate the CO seasonal cycle, although with a low-biased magnitude. The difference between observations and MACC reanalysis (bottom panel of Fig. 3.2) shows that larger differences are mostly located in the lower atmosphere.

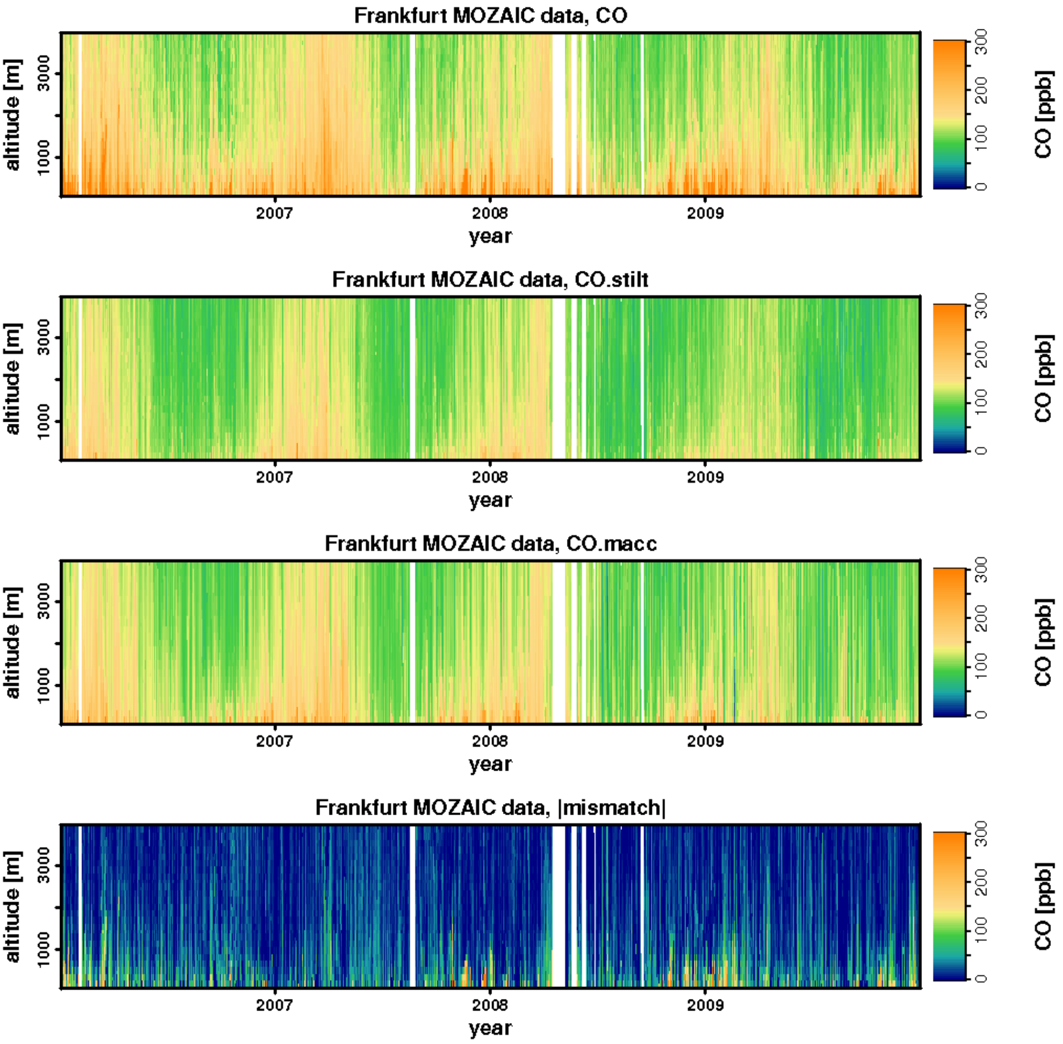

Figure 3.2: Observed MOZAIC/IAGOS profiles of CO near Frankfurt (first panel, on top), together with simulated profiles of CO from STILT/EDGAR coupling and MACC reanalysis (second and third panel respectively). The bottom panel shows the absolute value of the residuals between MOZAIC/IAGOS and MACC profiles.

As both models returned a net underestimate of the observed mixing ratio, an attempt to evaluate their ability to simulate observations was conducted. Figure 3.3 shows the coefficient of determination (R^2) between modeled and observed CO mixing ratio for the different intervals of the vertical profiles derived for both STILT/EDGAR and MACC. It is evident that the models' performance is very dependent on height and that they both perform similarly within the crucial area of the boundary layer, characterized by strong variability in the vertical transport.

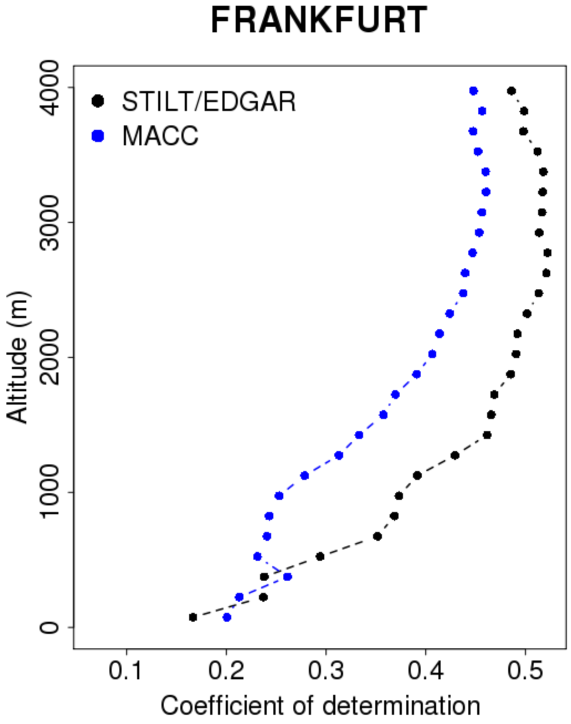

Figure 3.3: Coefficient of determination (R^2) between modeled and observed CO mixing ratio for both STILT/EDGAR and MACC using profiles collected around Frankfurt's airport in years from 2006 to 2011.

3.2 Corrections for mismatch in mixed layer heights

Comparisons between observed and model-derived heights of the mixed layer (z_i) show that the modeled values are biased low for both STILT/EDGAR and MACC, although for the latter the mismatch is lower (Fig. 3.4 shows an example limited to one year of Frankfurt data). More precisely, z_i modeled by STILT/EDGAR underestimates the observed value by 43% in Frankfurt, 53% in London and 29% in Vienna, while for MACC these values are 37% for Frankfurt, 46% for London and 17% for Vienna, respectively. From this low bias in simulated z_i one would expect a high bias in CO enhancements, corresponding to a correction that lowers the original CO enhancement. These values translate into average corrections for the CO enhancement simulated by STILT/EDGAR of roughly -6.8% (Frankfurt), -2.8% (London) and -9.3% (Vienna). Note that when limiting the statistics to only include those cases where modeled z_i is above 225 m (see 2.2.3), these values change to -29.9%, -15% and -29.3%, respectively. The corresponding values for the MACC reanalysis are 0.3% (Frankfurt), 6.5% (London), and 5.1% (Vienna) if all the dataset is considered, and -1%, 17.8%, and 8.5% if the 225 m filter is applied.

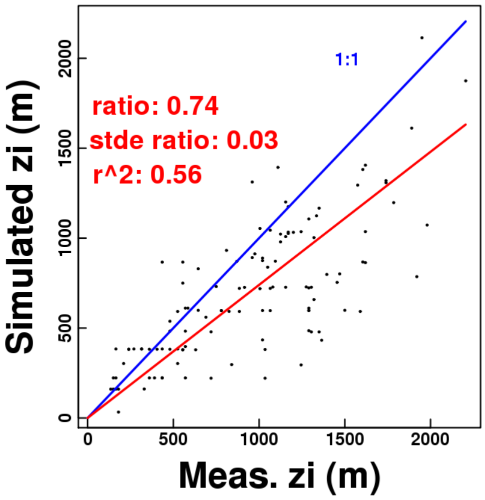

Figure 3.4: Comparison of simulated vs. observation-derived mixing heights for MOZAIC profiles near Frankfurt in 2007. The red line is drawn from the Origin and through the center of mass of the scatter plot, so its slope represents the ratio of the mean simulated and observed value.

3.3 Corrections for flux error

It is clear from Fig. 3.5 (left) that the CO enhancements from our STILT/EDGAR simulations are biased low. From the optimization of the modeled fluxes using equation 2.7, we found that both STILT/EDGAR simulated emissions of CO for both STILT/EDGAR and MACC reanalysis are biased low by a factor of 2-2.5 (Fig. 3.5, right). There is no significant difference in the scaling factors for STILT/EDGAR and MACC, as there is an overlap in the error bars of each specific correction factor. As error bars for the scaling factors of different cities also overlap, we can claim that there is no significant spatial correction. However, it is worth to point out that a slightly significant seasonal correction is indeed present, as error bars for the scaling factors of months belonging to different seasons do not always overlap. Values for different months range from about 0.8 during summer to 1.3 during winter.

The relevance of the seasonal pattern in the bias of emission models has been found to be important by Stein et al. (2014). However, their scaling factors for Europe range between 1 and 4.5 for monthly values (Table 4 from Stein et al., 2014), which corresponds to a range from 0.4 to 2.0 when separating out the overall scaling factor of 2.3, showing much stronger temporal variability than in the present study. In this study, temporal emission factors for different months, days of the week and hour of the day were applied to EDGAR annual fluxes specific for both emission category and fuel. Conversely, MACCity emissions datasets used by Stein (2014) are developed on a decadal basis, with a linear interpolation applied to obtain yearly fluxes. A source-specific seasonality was then implemented.

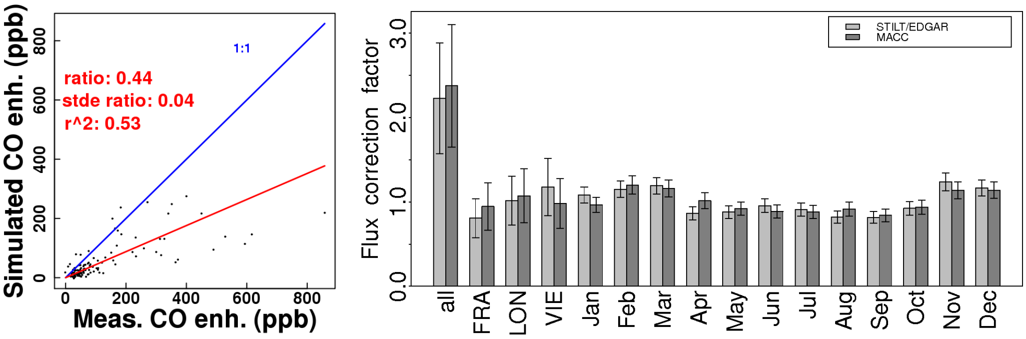

Figure 3.5: In the left panel the comparison of simulated vs. observed mixed layer CO enhancements for Frankfurt profiles in 2007 during daytime (10:30-17:30 UTC) is shown. The red line is drawn from the Origin and through the center of mass of the scatter plot, so its slope represents the ratio of the mean simulated and observed value. The right panel shows the correction factors to compensate for a bias in STILT/EDGAR and MACC emission flux (right). Both correction factors and error bars (standard deviations) were derived using a weighted least-squares estimate of the parameters of a non-linear model.

3.4 Evaluation of simulated CO enhancements

Before quantitatively evaluating the representation error, we evaluated the statistical model with respect to ML enhancements. We are especially interested in how well the STILT/EDGAR simulations and the MACC reanalysis can reproduce the observed enhancements before and after being corrected for bias in simulated fluxes. Being a high-resolution transport/flux coupling, STILT/EDGAR is expected to detect the near-field influence of emissions on the tracer enhancement.

Observed values for CO enhancements range between 30 and 130 ppb (Fig. 3.6) and are strongly dependent on wind direction. For all three cities, maximum values are observed when the wind blows from East, whereas minimum values are usually observed when wind blows from West. Observed enhancements over Frankfurt experience a maximum when wind direction is 45-75 deg., and a minimum when wind direction is in the interval of 255-315 deg. For London, maximum and minimum values are observed when wind direction is 75-105 deg. and 225 deg. respectively. For Vienna data are not available for many wind sectors; among the available sectors, maximum values are collected when wind direction is 105-165 deg., whereas minimum values are collected when wind direction is 285 deg. For both Frankfurt and London, maximum enhancements correspond to situations when measurements are recorded downwind the main source region (the city center). In addition for Frankfurt, where data coverage is better, enhancements observed under low wind conditions tend to be similar to the maximum values observed for the other 30 degrees sectors. This suggests that the emissions from the city center and from airport itself have the same potential to influence observations.

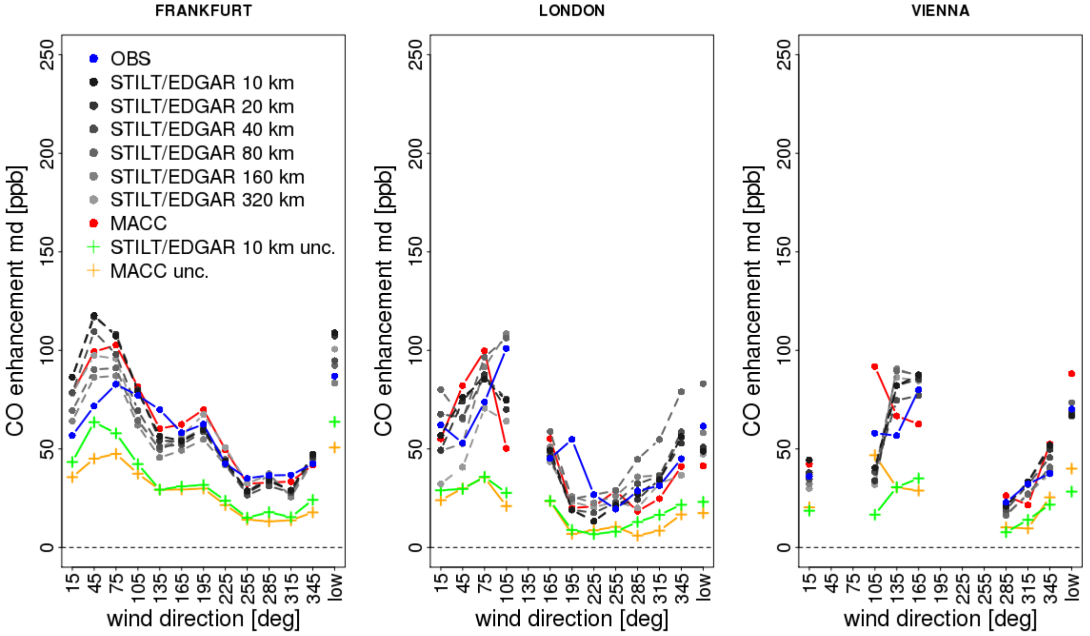

Figure 3.6: Median enhancements of CO for the years 2006-2011 in the mixed layer for Frankfurt (left), London (middle), and Vienna (right), as a function of wind direction. The rightmost x-values indicated "low" represent low wind speeds (<3m/s). Observations are show in blue, STILT/EDGAR simulations in different grey tones (light for coarse, dark for high resolution), and MACC reanalysis results are shown in red. STILT/EDGAR and MACC uncorrected enhancements are shown in green and ochre respectively.

It is clear from the same plot that both STILT/EDGAR and MACC need to be corrected to avoid a severe underestimation of the observed enhancement values. As for the corrected modeled outputs, they range between 30 ppb and 130 ppb (median over the years 2006-2011). In addition, the STILT/EDGAR simulation usually performs similarly to the MACC reanalysis in reproducing the observed enhancements in their dependence from wind direction. However, interpretation is not trivial, as the flight path depends strongly on wind direction (aircraft typically take off into the wind). It is worth to note that the enhancements derived from STILT/EDGAR for different spatial resolutions usually share a similar pattern, with a relative difference in magnitude up to 30% for Frankfurt, 50% for London and 20% for Vienna.

Note also that corrected STILT/EDGAR simulations at lower resolutions occasionally show better agreement with observations than their counterpart derived with more highly resolved fluxes. In fact, low-resolution CO enhancements tend to be reduced due to horizontal averaging of strongly localized sources; as uncorrected models systematically underestimate observed CO enhancement values, low-resolution uncorrected simulations will agree even less with observed values. However, after correction for flux error, STILT/EDGAR simulations can either overestimate or underestimate observations, which leads to the abovementioned effect.

Standard deviation results are shown in Fig. 3.7. It is found that corrected models have a higher standard deviation than the observations whereas uncorrected models have lower standard deviations. MACC and STILT/EDGAR at highest spatial resolution (10 km) have similar standard deviations, while STILT/EDGAR's standard deviation decreases together with spatial resolution.

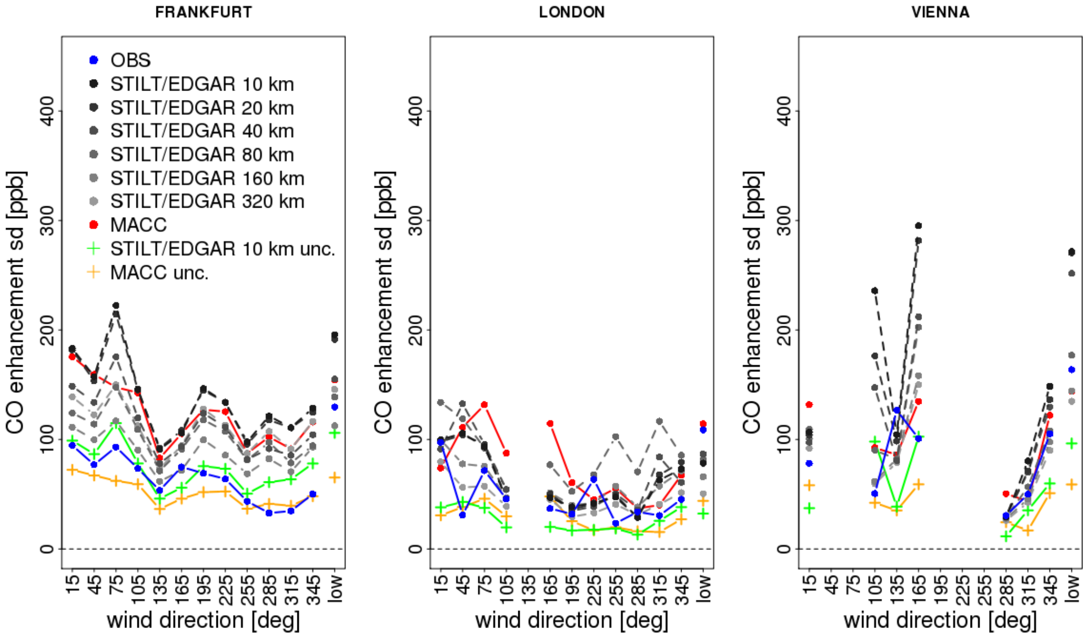

Figure 3.7: As Figure 3.6, but for the standard deviation of the enhancements of CO for the different wind sectors.

3.5 Representation error realizations

After applying the corrections for ML depth and flux error to the simulations, we plotted the realizations of representation error against the low-resolution enhancements in the ML (Fig. 3.8). Plots were done seasonally or by location; the spatial scale of 80 km was chosen as the closest to the MACC horizontal resolution (1.125 deg). It is clear from the plot that higher enhancements will lead to higher realizations for the representation error (see 5-95%ile envelopes as grey lines in Fig. 3.8). In addition to a larger variance, the mean of the error realizations for different simulated enhancements also shows a slight decrease for larger enhancements. In other words, the high-resolution simulations result on average in larger enhancements compared to coarser resolutions. This is related to the fact that local emissions near the IAGOS/MOZAIC observations made within the mixed layer are strong and extend over small areas, whereas they become more diluted when using coarser resolution. Results do not show any clear dependence on the season, but slight differences can be seen for different airports.

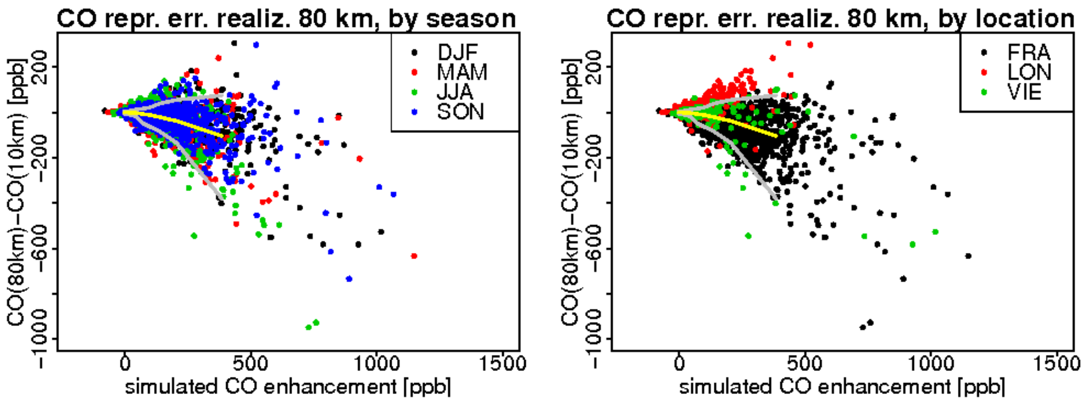

Figure 3.8: Realizations of representation error (i.e. differences between STILT simulations at different resolutions, here 10 km and 80 km) for CO plotted against simulated enhancement, and color-coded by season (left) and by airport location (right). Grey lines indicate the 5 and 95%ile of the distribution within 10 bins of simulated enhancement; the yellow line indicates the mean.

3.6 Representation error

Both random and bias component of the representation error are highly variable with spatial resolution (Figs. 3.9 and 3.10). Representation error tends to increase with decreasing resolution even though the general dependence on wind direction is conserved. Comparing the absolute representation error associated with the highest spatial resolution (20 km) with the representation error associated with the lowest spatial resolution (320 km), we found that such increase can be by a factor of 4-5 for the random component, and more than 10 for the bias component. For the random component, the representation error is around 2-10% for 20 km, and with few exception increases to around 10-100% for 320 km (Fig. 3.9), while the range for the bias component, the representation error is from -2% to +1% at the highest resolution, and from -50% to +50% at the lowest resolution (not shown).

Most of the remarks for the relative representation error also hold for the absolute representation error, especially the strong dependence on spatial resolution. For the random component, representation error is around 2-8 ppb for 20 km and increases to around 10-50 ppb for 320 km (Fig. 3.10). For the bias component, values can be negative; the representation error ranges from -3 to 1 ppb at highest resolution and from -40 to 40 ppb for the lowest resolution (not shown).

Note that the random component of the representation error increases from 160 to 320 km spatial resolution for London and Vienna, but decreases for Frankfurt; this effect is probably a result of specific property of the emission pattern around the Frankfurt airport, with local emissions (which have a strong influence on the CO enhancement) being more comparable to emissions aggregated to large scale than to those aggregated to intermediate scales.

The strong dependence of representation error on both spatial resolution and wind direction indicates that coarser models are expected to have difficulties representing the small spatial scale of the emissions around strong localized sources, for example those originating in the cities. This is most likely due to the effect of horizontal dilution that such averaging has on the emissions.

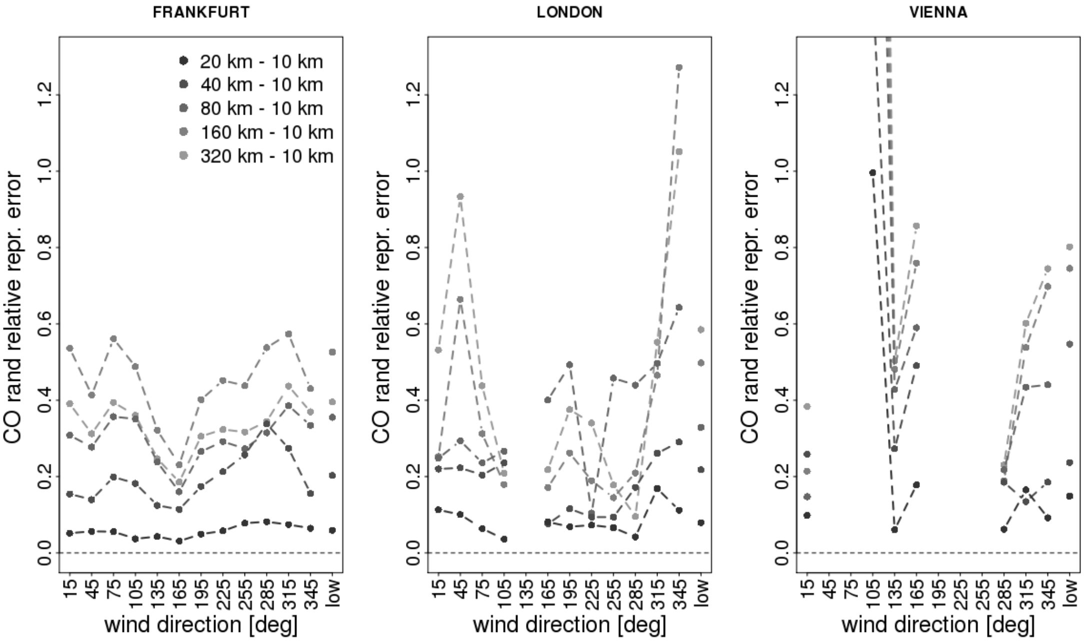

Figure 3.9: Random component of the relative representation error for CO for the years 2006-2011 in the mixed layer for Frankfurt (left), London (middle), and Vienna (right), as a function of wind direction. The rightmost x-values indicated "low" represent low wind speeds (<3m/s). STILT/EDGAR simulations are shown in different grey tones (light for coarse, dark for high resolution). Maximum relative error for Vienna at 105 degrees is up to 4.8

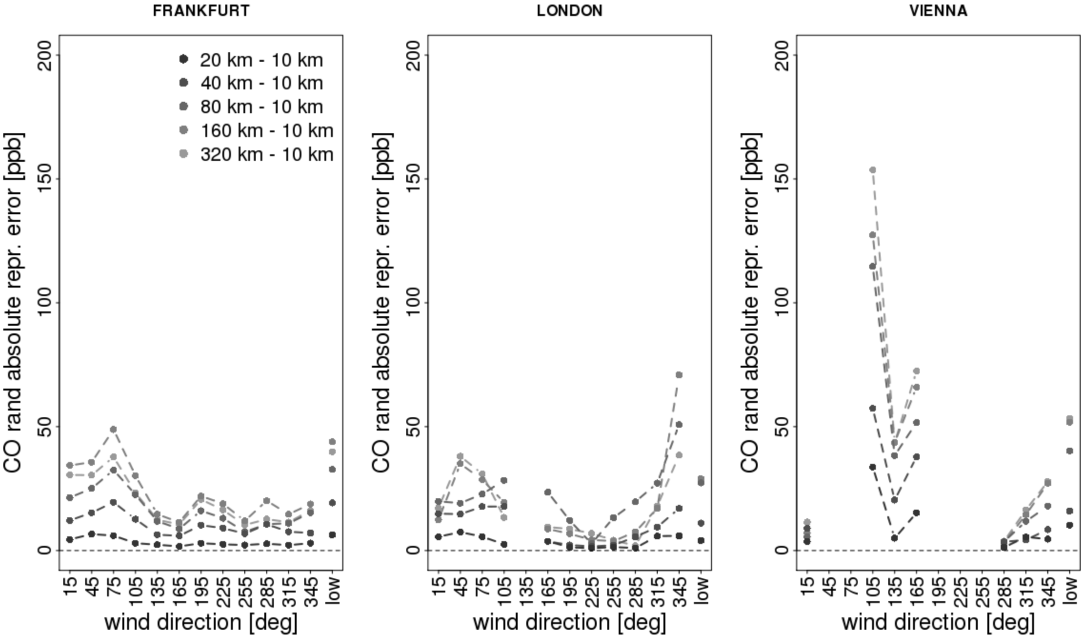

Figure 3.10: As Figure 3.9, but for absolute representation error.

3.7 Validation

In order to evaluate to which degree the representation error can be useful to describe and ideally predict model-data mismatch for an independent model, the representation error derived from STILT/EDGAR was compared with residuals between the MACC reanalysis and observations from MOZAIC/IAGOS. Here we assess the dependence on wind direction and on time (month) in order to evaluate whether or not the representation error has any capability to describe spatial or temporal variations in model-data mismatch (Fig. 3.11). The analysis is limited to the city of Frankfurt due to the better data coverage. The slope of the linear regression indicates the ratio of variance in the model-data mismatch that is accounted for by representation error and was derived using the Theil-Sen estimator. Such method calculates the median of all the slopes of the lines passing through a couple of points in the graph and is therefore less sensitive to outliers.

As mentioned in section 2.2.6, enhancements are expected to be more or less bias-free after flux error correction. Hence the bias component for representation error cannot be validated and for this reason, only results for the random component are shown. In Fig. 3.11 is shown that random representation error allows for describing 15-21% of the variance with good correlation coefficient returned for both wind and temporal grouping (0.58 and 0.67 respectively). Model-data mismatch ranges roughly over 60 and 180 ppb. This result suggests that the representation error provided by STILT/EDGAR can explain a significant fraction of the random component in the model-data mismatch for MACC, and therefore can be regarded as useful information for better understanding causes for model-data mismatch in other models.

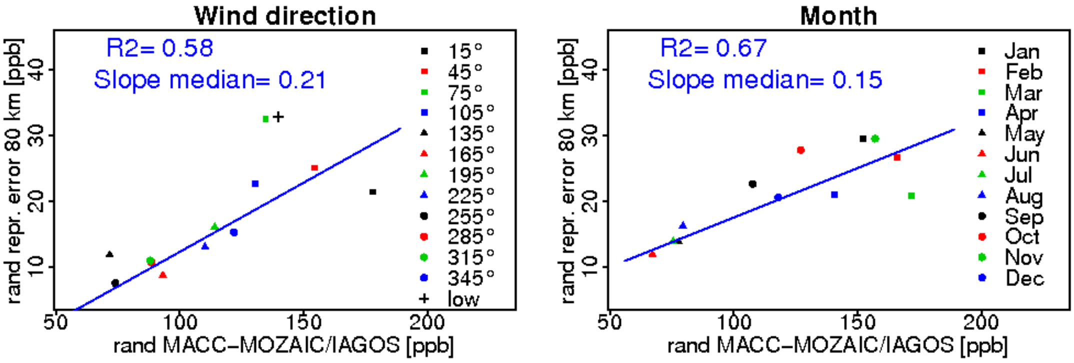

Figure 3.11: Random component for representation error of Frankfurt for different wind directions (left) and months (right), plotted against the corresponding model-data mismatch error.

3.8 Error contributions

After the individual errors related to mismatch in mixing height, bias in emission fluxes, and spatial representation have been quantified, they can be compared to determine their relative importance. As mentioned before (section 2.2.3), due to uncertainty in observations, we apply the correction for mixed layer depth only when the observed z_i is higher than 225m, which means in 55% of the cases. Note that of the remaining cases, only 35% for STILT/EDGAR and 25% of MACC occur during daytime (11:00 – 17:00), and the rest during nighttime or transition periods. This means that for almost half of the data the z_i correction factor equals one, which from the model's perspective is equal as saying that they don't need to be corrected although this is certainly not the case. To account for this effect, when we evaluate the error contributions from different categories of uncertainty, we perform our investigation only on the sub-dataset in which both the mixed layer mismatch and flux correction are implemented.

The assessment of contribution of different error categories to model-data mismatch for the city of Frankfurt is shown in Fig. 3.12. Here the random and bias components are treated separately. Random and bias component for mixing layer mismatch and flux correction are here calculated according to Equations 2.9 and 2.10, whereas both components for the spatial representation error are calculated as the mean of respective component derived in section 2.2.5. The fact that the second source of uncertainty (flux error) has also a random component is related to the fact that simulated fluxes are also corrected with a time dependent (monthly) factor. Thus when comparing simulated CO enhancements before and after correction for flux error (overall and monthly), the standard deviation of the introduced relative differences reflects the temporally varying flux corrections. Note that the different contributions are calculated as relative errors for each category, not as the fractional contribution to the total error (i.e. the sum of the contributions is not necessarily equal to one).

Random components for the mixing layer mismatch are about 83% for STILT/EDGAR simulation and 70% for MACC. The random component for flux correction is around 8% for STILT/EDGAR and 5% for MACC. Contributions from representation error range from 14% to 89% according to the considered resolution. The bias component of model-data mismatch is positive for z_i mismatch and negative for flux correction. More detailed, the relative error for mixed layer depth is 55% for STILT/EDGAR and 26% for MACC, while for inaccuracies in simulated fluxes such values are -44% and -55% respectively. Note that both random and bias component of the representation error increase with decreasing resolution up to 160 km, a feature observed also in the Frankfurt panel of both Fig. 3.9 and Fig. 3.10.

Further sources of error with relatively small impact are photochemistry and uncertainty due to boundary conditions. As stated in section 2.2.2, the contribution to modeled CO from the advected lateral boundary condition was subject to photochemistry (reaction with OH and production from CH_4 oxidation), whereas no chemical loss was assumed for the additional CO from emissions within the domain. In fact, observations tend to be influenced the most from local sources, with the median of the time (prior to measurement) needed to account for 90% of the CO contribution from emissions is 15 hours, and with half of the CO contribution to each trajectory ensemble captured in 25 hours for 90% of the measurements. According to Protonatariou et. al (2010), CO lifetime during summer in Europe ranges from weeks to month. Using 14 days as an extreme value for CO lifetime, the amount of CO oxidized by OH in 15 hours would be to roughly 4.4% of the enhancement. We regard this as the upper limit of the uncertainty (both bias and random) introduced by neglecting photochemistry for CO from emissions within the domain.

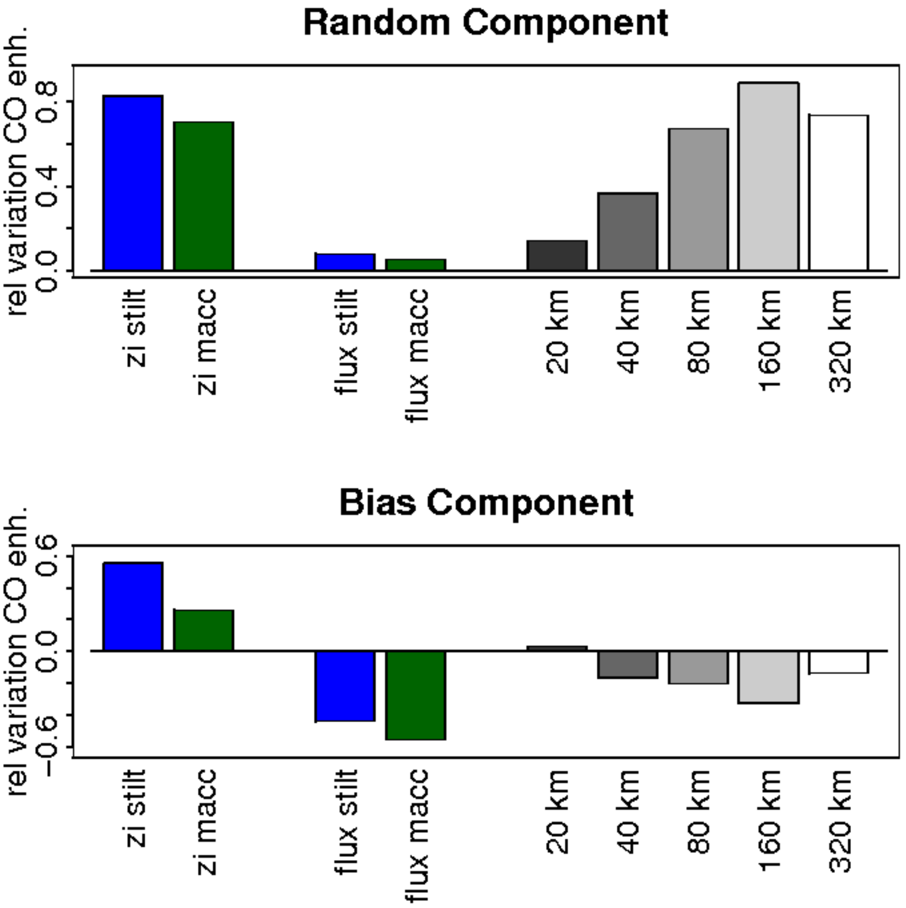

Figure 3.12: Assessment of contribution of different error categories for the city of Frankfurt. The assessment is treated separately for random (upper tab) and bias (lower tab) component. For each component the uncertainty of the correction for mismatch in the mixing height (z_i) and bias in the emission inventories (flux) is shown for both STILT (left) and MACC (center) models; the contribution from the spatial resolution of EDGAR fluxes to STILT/EDGAR uncertainty is shown for each of the considered resolutions (right).

The change in the simulated CO enhancements due to photochemistry (taken as the difference between simulations with and without photochemistry) are shown in Fig. 3.13 (left); the mean and standard deviation of these changes amounts to -0.7 and 2.6 ppb respectively. Note that the photochemistry not only accounts for losses in CO, but also for CO generated by the oxidation of methane. We regard these differences as upper limit for the bias and random error resulting from imperfection in the chemistry as applied to the lateral boundary condition. Together with the uncertainty due to neglecting photochemistry for the CO emitted within the domain, the overall uncertainty in photochemistry is estimated to be on the order of 5.5% (random) and 8.9% (bias) for the CO enhancement.

For the uncertainty due to boundary conditions we follow a similar approach by calculating residuals between STILT/EDGAR with MACC boundary condition and corresponding simulations with zero boundary condition (Figure 3.13, right). The mean of the residuals is 2.9 ppb, while the standard deviation is 11.1 ppb, corresponding to 4% and 17.1% in relative terms, given average enhancements of 64 ppb.. Again we regard these as upper limits for the corresponding uncertainty due to imperfect lateral boundary conditions.

Random contributions to uncertainty in the CO enhancements from both photochemistry and boundary conditions (5.5% and 4%) are comparable to those from flux correction, but much smaller than contributions from the mixed layer depth and representation error. Note that the random error contribution from flux error is only related to the monthly variations in the correction factors, which are relatively small. As for the bias contributions from photochemistry and boundary conditions (8.9% and 17.1%), they are small compared to the other error contributions.

Other sources of error not explicitly taken into account here include uncertainty in horizontal transport and the occurrence of deep convection due to the presence of clouds. Uncertainty in simulated transport related to poor modeling of advection was thoroughly described by Lin and Gerbig (2005), where the (relatively favorable) comparison between modeled wind from the Eta Data Assimilation System (EDAS) with observation from radiosondes was used to specify uncertainty in simulated winds. Unfortunately the MOZAIC/IAGOS wind observations do not show a nearly as good agreement with the ECMWF simulated winds, most likely due to observational difficulties for airborne platforms as compared to radiosondes. We thus refrain from attempting to quantify this source of uncertainty. Uncertainty related to imperfect representation of vertical transport by deep convection is not likely to contribute strongly, as it occurs relatively infrequent.

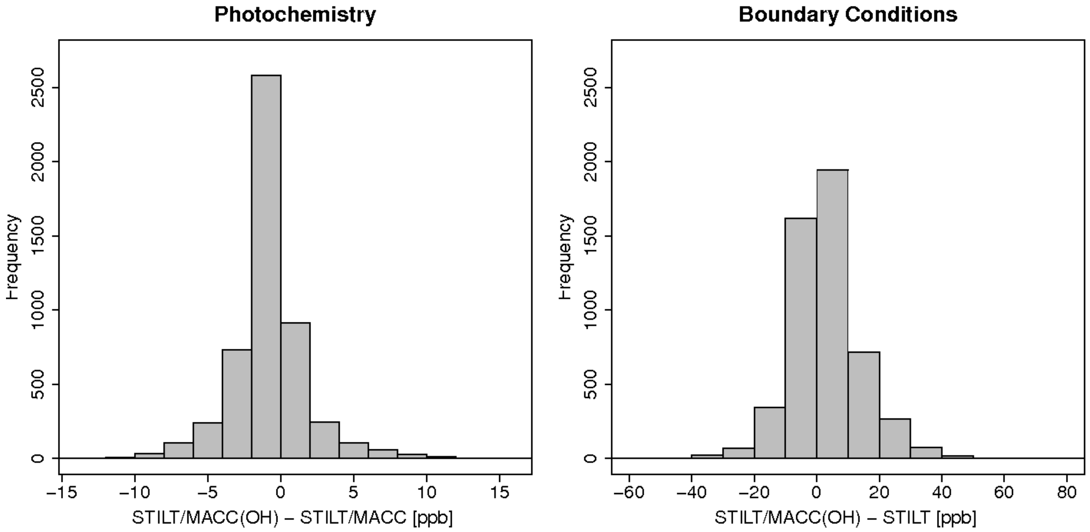

Figure 3.13: Absolute change in CO enhancements due to photochemistry (left) and boundary condition (right) on the whole dataset. Standard deviation of residuals is quantified as 2.6 ppb for photochemistry and 11.1 ppb for the boundary condition. Note the different scale on the x-axis.

Chapter 4 : Multi-species inversion and airborne data for a better constraint of continental scale fluxes

After describing the different sources of error while modeling MOZAIC/IAGOS airborne data in Chapter 3, this chapter is centered on improving the constraint of atmospheric inversion when using airborne data. In particular, we assess the benefit of explicitly consider the correlations between different species in order to evaluate the effect of upcoming simultaneous retrievals of CO_2, CO and CH_4 in the frame of the IAGOS project.

4.1 Observed and modeled mixing ratios

Before evaluating the performance of the inversion scheme in reducing the uncertainty of the state space, a closer look at the ability of the modeling framework to reproduce the enhancements is necessary. Unfortunately, this can be done only for CO as actual measurements are not available for the other species. Figure 4.1 shows the mean daily enhancement of the three fossil fuel species for both model outputs and observations. A common feature to the three trace gases is that lower values tend to occur during summer time due to a better mixing of the atmosphere. Conversely, enhancement values tend to be higher during winter, reflecting the more stratified atmosphere of the cold months.

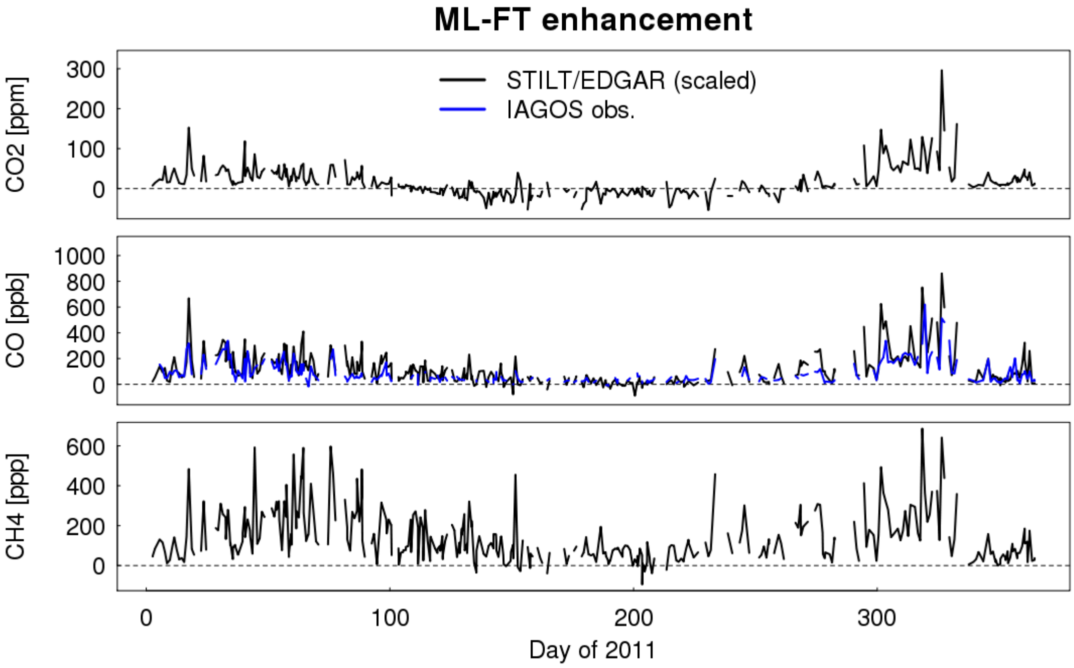

Figure 4.1: Mean daily enhancement of mixed layer vs. free tropospheric mole fractions. Modeled mixing ratios are shown as black lines, while the observed CO is shown as blue line. Note that the modeled values for CO have been multiplied by a factor of 3.8, corresponding to the mean ratio between observed and modeled CO enhancements (after z_i correction), to match the observed values.

In Fig. 4.1 the modeled CO plot was multiplied by a factor of 3.8, corresponding to the mean ratio between observed and modeled CO enhancements, similar to what was found in Chapter 3 for the year 2011. Mixing ratio values are highly variable, but the model usually manages to reproduce the associated spikes; the squared correlation coefficient between observed and modeled CO enhancements is 0.62 and the standard deviation of corrected model and observation residuals is 85 ppb. The median of the mixing ratio enhancement for the three trace gases is 18.6 ppb for CO, 26.6 ppb for CH_4 and 2.8 ppm for CO_2. For CO_2 this seasonal difference is enhanced due to the simultaneous presence of both anthropogenic and biogenic emissions. During summer values are slightly negative due to strong photosynthesis fluxes from growing vegetation combined with deeper vertical mixing. Negative values arise in 31% of the cases predominantly during the warmer months, implying that during the growing period uptake by photosynthesis dominates over release from combustion and respiration. Both CO and CH_4 experience higher values during winter due to the shallow mixed layer usually associated with cold temperatures, and lower values during summer as higher temperature cause the mixed layer to reach higher altitudes; differences related to seasonal domestic heating and transportation may also play a role. In addition, enhancement for both species is occasionally negative, most likely owing to advection of polluted air masses in the free troposphere. An alternative explanation is that strong winds at lower heights can disperse the emissions in the boundary layer and create a situation in which the mixing ratio in the FT is higher than in the ML.

4.2 Multi-species inversion

With respect to the prior error covariance matrix, the posterior error covariance shows lower values (Fig. 4.2) corresponding to an uncertainty reduction of 22%.

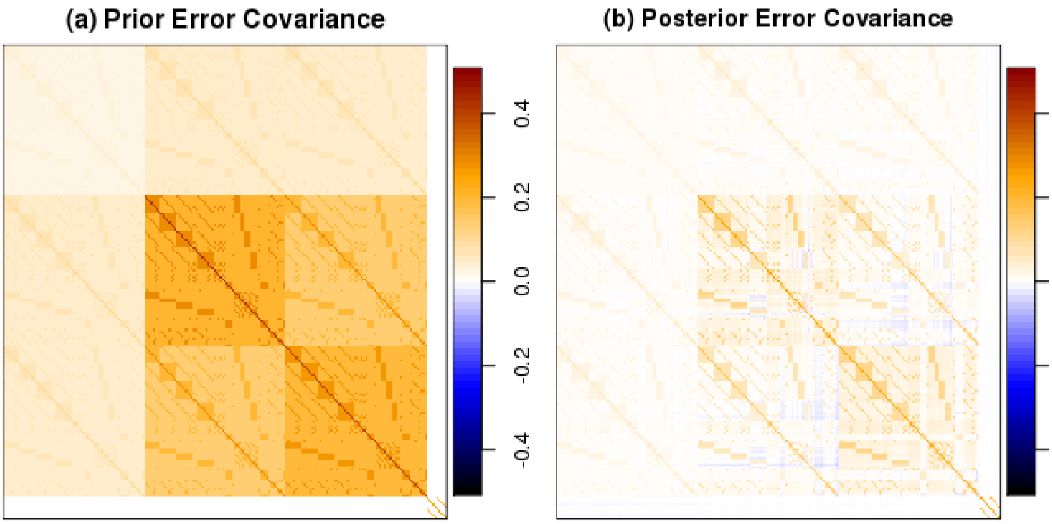

Figure 4.2: Prior error covariance matrix (left) and corresponding posterior error covariance matrix (right).

Figure 4.3 and 4.4 show a priori, a posteriori, and "true" fluxes related to different aggregated fuel types and to different emission categories as described in Tables 4.1 and 4.2 for the months of July and December. Figure 4.4 also shows the biospheric contribution (as absolute values) scaled down by a factor of 10. As it is to be expected, the biospheric contributions show strong differences according to the seasonal cycle, while anthropogenic emissions remain rather stable. However, it is worth pointing out that while the fossil fuel prior is similar for both months, the assumed truth can be rather different due the random assignment of the prior error realization. In most cases, the posterior adapts and is therefore closer to the truth than the prior; the posterior uncertainty is also visibly reduced, as expected. Regarding the different tracers, CO_2 and CO show a somewhat similar pattern indicating a partial overlap in dominating emission categories while CH_4 is dominated by different contributions in both fuel types and emission categories.

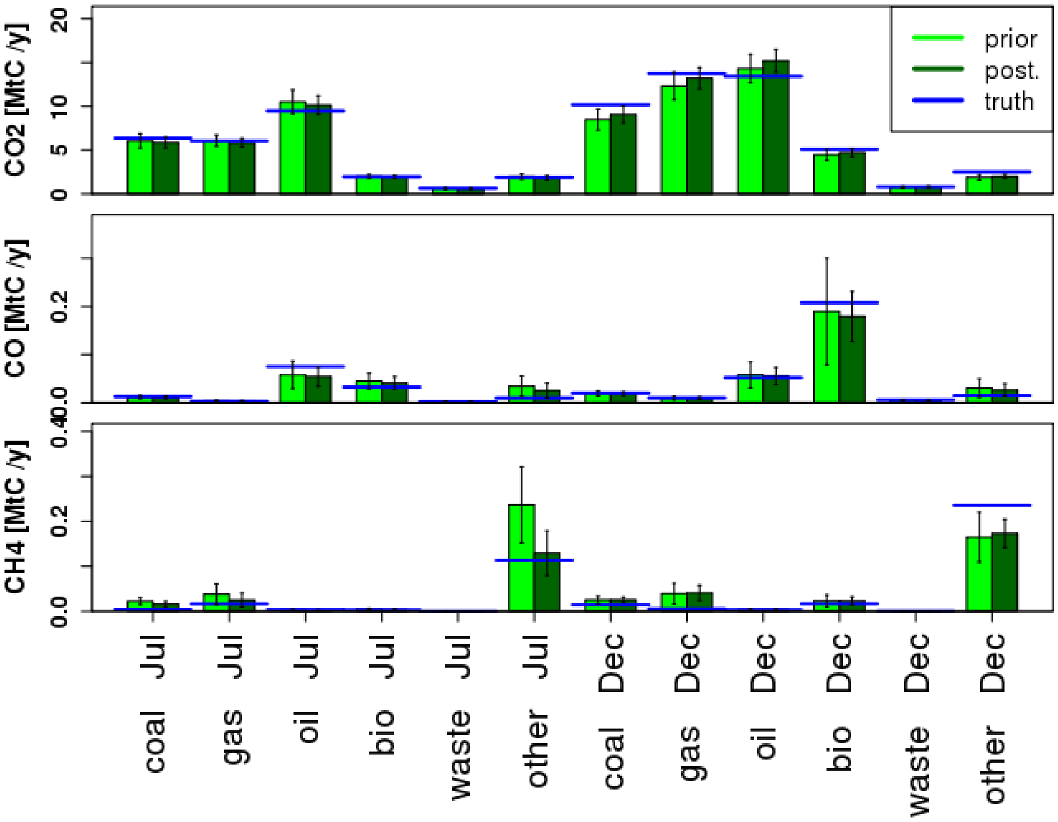

Figure 4.3: Prior, posterior and true (pseudo-data) fluxes in physical units aggregated for different fuel types. Note that, as the true fluxes are the result of a random perturbation of the prior, they do not describe an actual situation in the physical world. So, for example, the fact that the true value of CH_4 fluxes in July is lower than the same value in December should not be surprising.

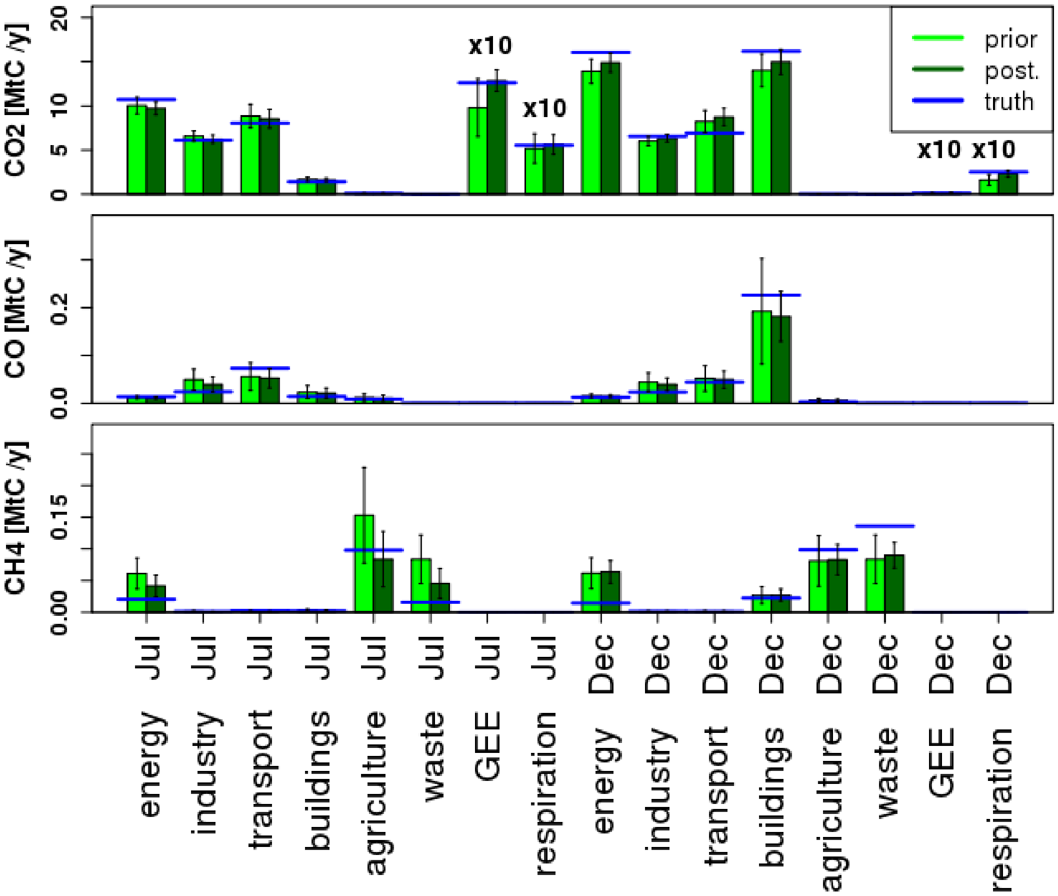

Figure 4.4: Prior, posterior and true (pseudo-data) fluxes in physical units aggregated for different emission sectors. Absolute values of biosphere-atmosphere exchange fluxes of CO_2 are included in (b), but scaled down by a factor of 10. Note that, as the true fluxes are the result of a random perturbation of the prior, they do not describe an actual situation in the physical world. So, for example, the fact that the true value of CO for transport in July is higher than the same value in December should not be surprising.

Dominant fuels for CO_2 are coal, gas and oil, whose prior fluxes (pseudo data) have a magnitude of 6-11 Megatons of carbon per year (MtC y^{-1}) in July and 8-14 MtC y^{-1} in December, while CO is dominated by a 0.19 MtC y^{-1} flux from biofuels during winter and secondary contributions during summer from oil and biofuels with a magnitude of 0.06-0.08 MtC y^{-1}. Regarding CH_4, the single dominant contribution is from "Other" fuels, responsible for 0.16-0.24 MtC y^{-1} of emissions. "Other" fuel types include emissions from non-metallic minerals industry (e.g. cement, lime), agricultural waste burning, metal industry processing, chemical and solvent industry, solid waste disposal in landfills, wastewater treatment, manure management in agriculture, rice cultivation in agriculture, and agricultural soil emissions. For CO_2, the dominant contribution from these "Other" fuels in the European domain is from the non-metallic mineral industry (1.13 MtC y^{-1}); for CO and CH_4, the lion's share of the "Other" fuels emission is from the metal industry (0.03 MtC y^{-1}) and agricultural waste burning (0.06-0.11 MtC y^{-1}).

The most important emission sectors for CO_2 are energy, industry, transport and building, each contributing 7-10 MtC y^{-1} in July and 6-14 MtC y^{-1} in December, while CO is dominated by a 0.19 MtC y^{-1} flux from buildings during winter with secondary contributions from industry and transport with a magnitude of 0.04 MtC y^{-1} and 0.05 MtC y^{-1} respectively in both the analyzed months. CH_4 is dominated by a contribution of 0.15 MtC y^{-1} flux from agriculture in July with secondary contributions from waste and energy with a magnitude of roughly 0.06-0.08 MtC y^{-1} in both July and December. The contribution from biospheric primary production is about 100 MtC y^{-1} in July, which drops to almost zero in December, while respiration values are 50 MtC y^{-1} in July and roughly 150 MtC y^{-1} in December.

As further assessment of the inversion performance, we tested the ability of the inversion scheme to capture the truth compared with a perturbed version of the prior. To do so we calculated for each simulated species the overall bias for the whole year between the prior and both posterior and the perturbed prior. It is clear from Table 4.1 that while the overall bias between posterior and truth is lower than the prior-truth bias, the bias between perturbed prior and truth is much higher, implying that the performance of the inversion is not an artifact of the pseudo-data generation.

	Prior - Truth (MtC y^{-1})	Posterior - Truth (MtC y^{-1})	Pert. Prior - Truth (MtC y^{-1})
CO_2 ff	79.2	67.5 (-15 %)	104.1 (+31 %)
CO	1.3	1.0 (-28 %)	1.9 (+42 %)
CH_4	2.2	1.0 (-52 %)	3.0 (+37 %)
GEE	184.1	79.3 (-57 %)	373.1 (+102 %)
Respiration	151.4	91.2 (-40 %)	221.2 (+46 %)

Table 4.1: Overall bias for different species between the prior and both posterior and perturbed prior. The percentage values in parenthesis refer to the corresponding Prior-Truth bias.

4.3 Inter-species correlation

Before investigating the benefits of correlations between different tracers, it is meaningful to evaluate the uncertainty reduction in the monthly budgets for all five modeled species (Fig. 4.5, based on targeted spatial domain in Fig. 2.8). The first thing to note is that for all of the five trace gases the posterior uncertainty is lower than the prior one, as it should be. In addition, prior uncertainty varies through the year, reflecting modulation in emission fluxes obtained by adding activity factors to describe the seasonal, weekly and daily cycle.

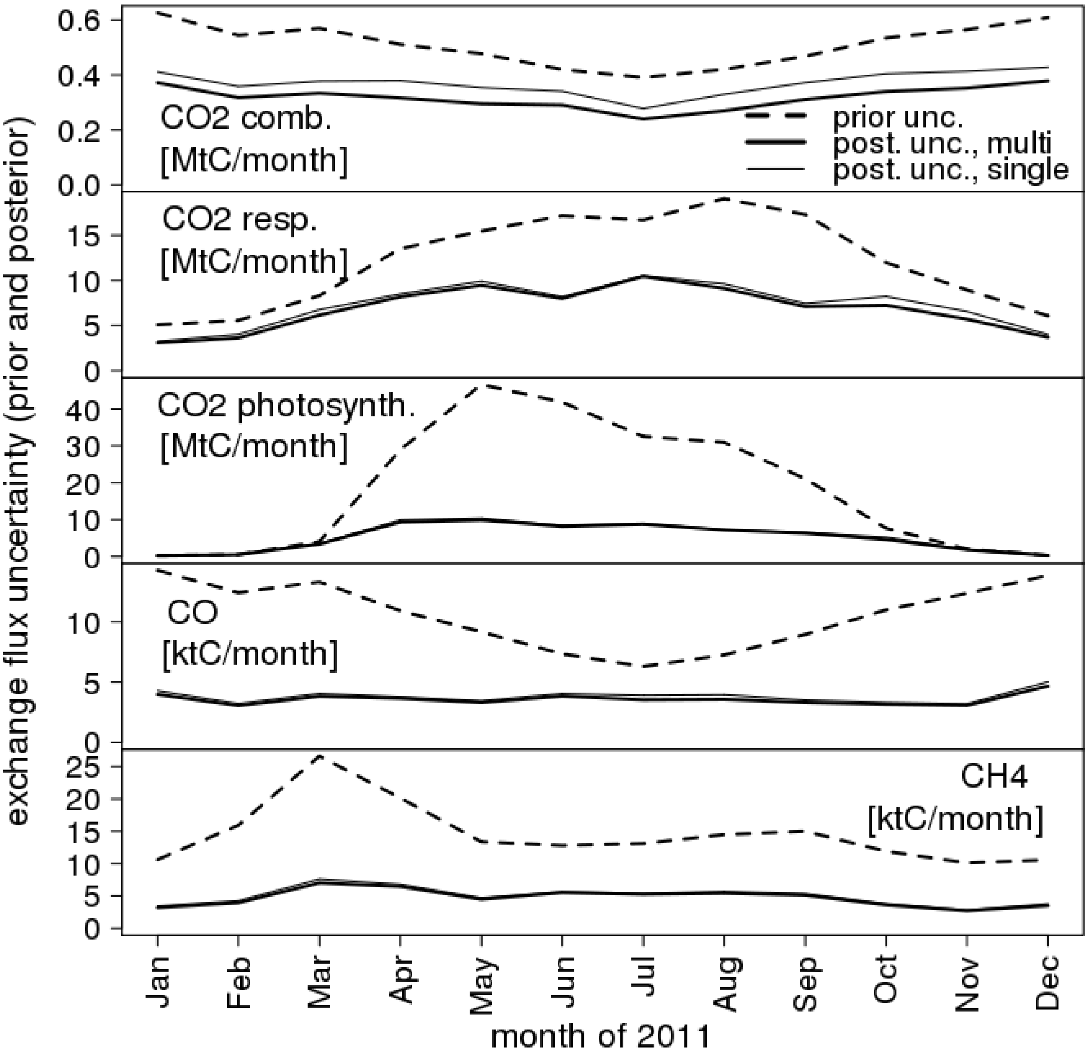

Figure 4.5: Comparison between prior and posterior monthly uncertainties for the five tracers. The posterior uncertainty is plotted for both the multi-species inversion, accounting for inter-species correlations, and the single-species inversion, in which all of the species are independent. Both prior and posterior uncertainty are expressed in physical units. The spike in the prior methane uncertainty estimate for the month of March depends on the emission inventory and is related to the cycle of agricultural activities.

Prior uncertainty assumes values around 0.4-0.6 MtC month^{-1} for CO_2, 5-15 ktC month^{-1} for CO, and 15 ktC month^{-1} for CH_4. For GEE the prior uncertainty is between 0.3 MtC month^{-1} and 46.7 MtC y^{-1}, and for respiration it is 5.1-19.0 MtC month^{-1}. Posterior uncertainty for CO_2 is 0.24-0.38 MtC month^{-1} for fossil fuel emissions, 0.3-9.9 MtC month^{-1} for GEE and 3.1-10.4 MtC month^{-1} for respiration, while it has a range of 3.3-4.7 ktC month^{-1} for CO and 2.7-7.0 ktC month^{-1} for CH_4. Mean uncertainty reduction of the monthly values is 38% for fossil fuels emission of CO_2, 41% for GEE, and roughly 45% for respiration, 64% for CO and about 67% for CH_4. It is worth pointing out that such values are higher than the mean uncertainty reduction in the scaling factors (22%); this happens because the most representative emission sectors are those influencing the observations the most and thus are also the most constrained.

In addition, note that in this case, the posterior uncertainties for single- and multi-species inversions are similar for the modeled species, with the exception of the CO_2 anthropogenic contributions. To generalize this last result, we tested the benefit of a multi-species inversion for the different cases of prior uncertainty values shown in Table 2.3. As an indicator for the benefit of including correlation between different species, we use the ratio between posterior uncertainty of the multi-species inversion and the posterior uncertainty of the corresponding single-species inversion. A value of one means that there is not benefit in adding an inter-species correlation to the inversion, while values greater than one means that a multi-species inversion is even less constrained than a single-species one. We expect this indicator to be less than one, meaning that inter-species correlations actually improve the constraint power of the inversion. As before, we consider here the uncertainties of the retrieved budgets for the 50% footprint, where the surface influence is strongest (Fig. 2.8). Values of this uncertainty ratio for the different trace gases as function of month are shown in Fig. 4.6 for the different cases listed in Table 2.3.

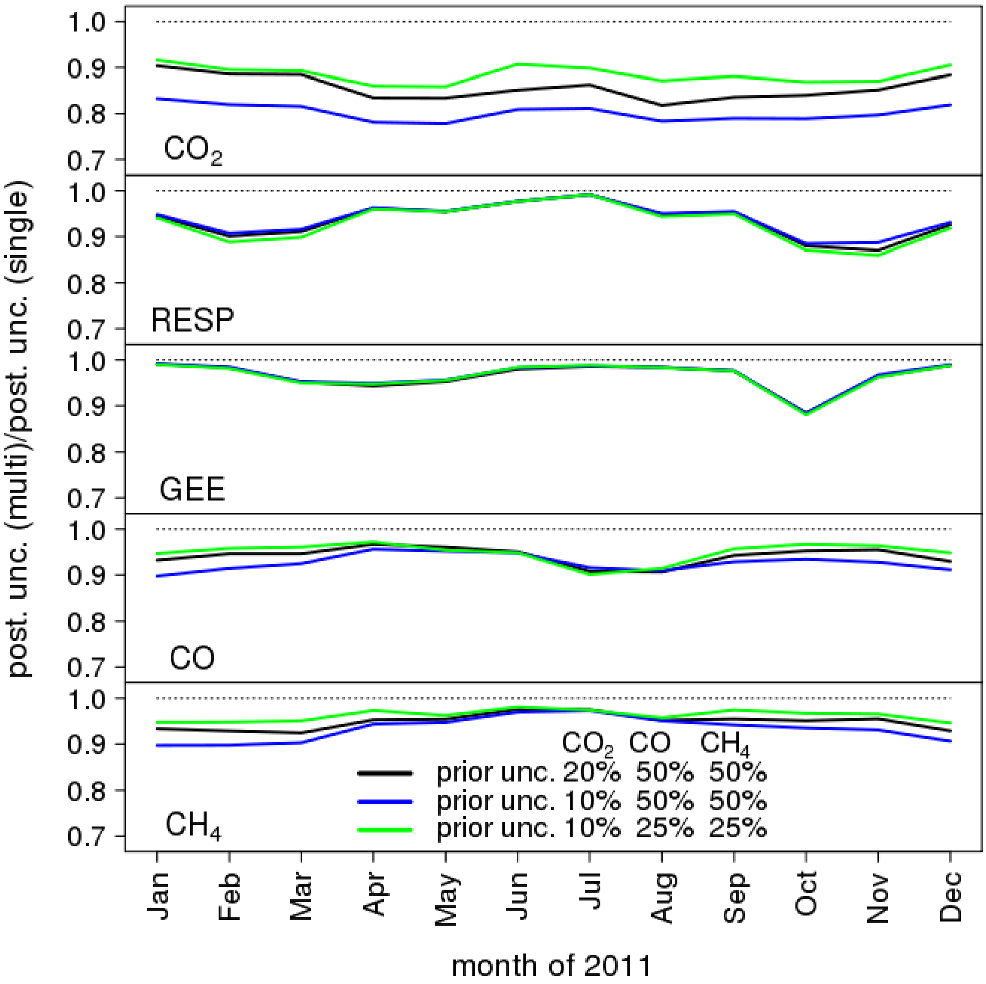

Figure 4.6: Benefit of a multi-species inversion over the corresponding single-species (dotted line) per different species per months of the year. The benefit has been tested for the three different cases of Table 2.3. Note that CO_2 refers to fossil fuel emissions only, and RESP and GEE refers to the biospheric fluxes.

All of the species experience a reduction in the posterior uncertainty ratio due to the addition of inter-species correlation; said reduction is up to 20% for fossil fuel CO_2 and up to 10% for the other species; In addition, anthropogenic CO_2 is more sensitive to the prior relative error values than CO and CH_4. As the uncertainty of GEE and respiration is not modified, they show little to no variations for different cases (Fig. 4.6). There is a dependence of the benefit of the multi- over a single-species inversion on the prior uncertainty values (differences between cases 1-3), with the largest difference for fossil fuel emissions of CO_2. Interestingly for case 2 with reduced prior uncertainty for fossil fuel CO_2 emissions the benefit nearly doubles over the default case (Case 1). Also reducing the prior uncertainties of CO and CH_4 emissions (Case 3) more or less compensates for this increase in benefit. Note that the assumed prior uncertainties for the default case (Case 1) are quite conservative, therefore lower uncertainties were chosen for Cases 2 and 3. While the absolute benefit of adding inter-species correlation is not a game-changer, it is worth pointing out that such improvement also comes with only slightly greater computational effort than multiple independent single-species inversions.

In order to assess the contribution of inter-species correlation in the prior uncertainty vs. that of model-data mismatch uncertainty, Fig. 4.7 also shows the resulting posterior uncertainty ratios for Case 1 (Table 2.3) from inversions only using prior or model-data mismatch correlation. For the anthropogenic component of CO_2, the greatest constraint is given by the prior correlation, while for GEE, respiration, and CH_4 the strongest contribution is from the model-data mismatch correlation. In the case of CO, the inter-species correlations for different components are dominant for different months of the year.

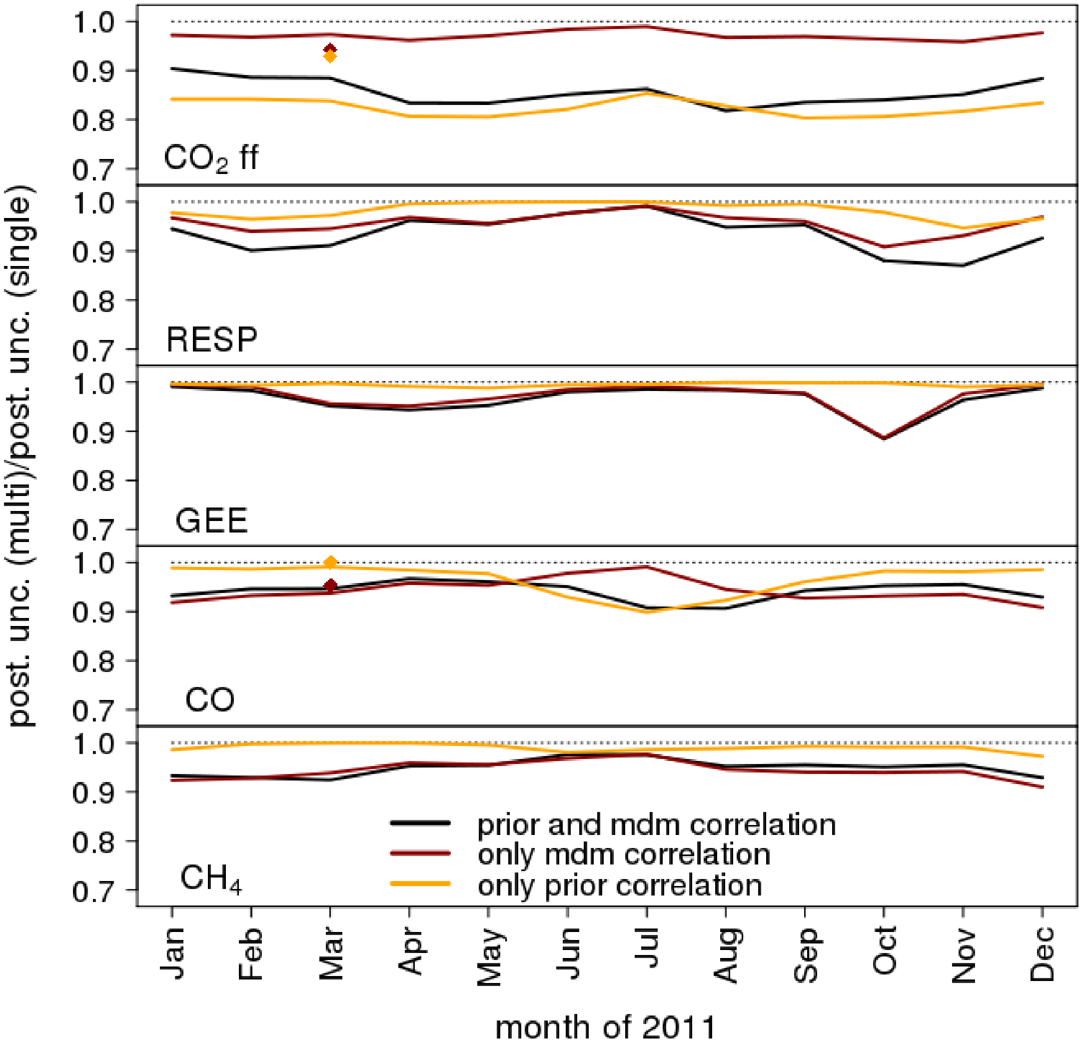

Figure 4.7: Benefit of a multi-species inversion over the corresponding single-species (dotted line) per different species and month. The benefit has been tested for a "normal" inversion featuring both prior and model-data mismatch correlation between different species (black) or only one of these two components (red and orange). Results refer to Case 1 of Table 3 (black line of Fig. 4.6). Values derived from Palmer (2006) for the month of March are indicated with a diamond.

Palmer (2006) (in the following referred to as P06) studied the importance of inter-species correlation to improve inverse analysis using airborne data from the TRACE-P mission conducted in March/April 2001 over the western region of the Pacific Ocean. P06 derived a prior error correlation lower than 0.2 by analysing the uncertainty of emission factors from an Asia-specific emission inventory (Streets, 2003), which is significantly smaller than the correlation of 0.7 assumed in the present study. P06 deemed CO_2-CO prior correlation greater than 0.5 to be unrealistic for the emissions in Asia, which is mostly associated with the uncertainty in emission factors for CO of 67% for fossil fuel emissions and 240% for biofuel emissions in China (P06 Table 1). However, for the European region used in the present paper we argue that values around 0.7 are appropriate. The resulting uncertainty in the CO_2-CO ratio, diagnosed from the prior error covariance matrix used in this study, is about 50% for both biofuel and fossil fuel emissions in Europe, which we regard as reasonable. To compare results from P06 with those from the present study, ratios of posterior uncertainties resulting from inversions using correlations between CO_2 and CO of 0.7 in the prior uncertainties and to those using no correlations have been extracted from Fig. 7 in P06 and are also shown as orange diamonds in Fig. 4.7. It is easy to see that for anthropogenic CO_2, the value derived from P06 is higher than in our study, while the two values are very similar for CO. Similarly, posterior uncertainty ratios using model-data mismatch correlations of 0.7 between CO_2 and CO are derived from Fig. 8 of P06 and are shown as red diamonds. In this case, the value derived from P06 is slightly lower than in our study for anthropogenic CO_2, while the two are again very similar for CO.

From this comparison we can see that the estimates of the benefit of including inter-species correlation in atmospheric inversions in P06 and in this paper are on the same order of magnitude for anthropogenic CO_2 and almost identical for CO, suggesting a general continuity of results.

Chapter 5: Using a transport model as a bridge to compare in-situ and column airborne observations

As explained in the Introduction, this chapter is focused on the newly developed CHARM-F lidar, carried by the HALO aircraft; being CHARM-F a new instrument, it still has to be tied to international standard from WMO. The investigation attempts to estimate the traceability of the lidar measurements back to WMO standard scales by making use of in-situ measurement from the JIG sensor, also carried by HALO, and a receptor-oriented modeling gramework (cfr. Section 1.4).

5.1 Model calibration

The results for the calibration of the STILT modelled mole fractions are summarized in Figure 5.1, where the mean of the 17 profiles is plotted for both CO_2 and CH_4 before (left) and after (right) bias correction. It is worth noting that boundary conditions tend to be coincident with the STILT profile in the Stratosphere and Free Troposphere, while the two strongly differ in the lower part of Troposphere, where surface influence is stronger. The mean profile clearly indicates a bias in the model-data mismatch in the Free Troposphere; after applying the daily correction profile to each of the modelled profiles, the model-data mismatch (MDM) is greatly reduced. More precisely, after correction the mean MDM for CO_2 is reduced by 91% from -1.1 ppm to -0.1 ppm, and the standard deviation of the MDM declines from 2.2 ppm to 1.8 ppm, corresponding to an 18% reduction. Similar results are found for CH_4, where mean MDM goes from 25.3 ppb to -0.4 ppb (98% reduction), while the MDM standard deviation declines from 27.2 ppb to 21.7 ppb (20 % reduction). By excluding profiles with particularly high MDM in deriving the correction profiles, similar results are obtained. This is probably due to the fact that these high-MDM profiles are characterized by both positive and negative MDM values that end up cancelling each other.

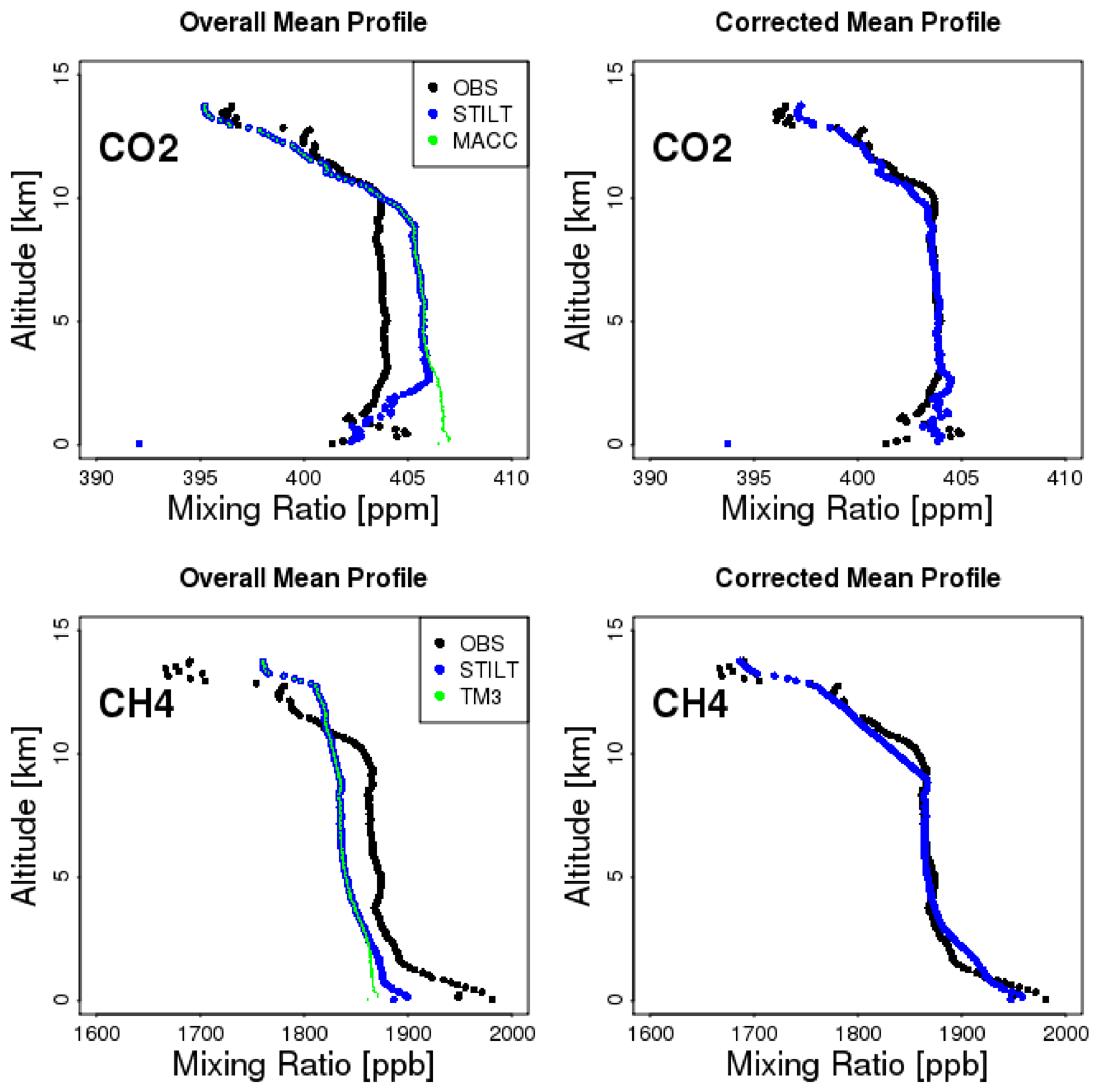

Figure 5.1: Comparison for CO_2 (top) and CH_4 (bottom) between mean observation (OBS, black dots), modeled profiles (STILT, blue dots) and advected boundary condition (MACC and TM3, green dots). The mean is calculated over all of the 17 profiles of the CHARM-F validation campaign. Modeled profiles are before (left) and after (right) bias correction.

5.2 Error propagation

After correcting the modeled profiles to match the observed profiles and converting them into slanted partial columns (SPC), we evaluated residuals between modeled and observed SPC and derived the respective SRE for the modeled SPC by calculating the standard deviation of five model output realization as described in Sect. 2.4.6. It was found that, the values of the model-data mismatch in proximity of the ground were disproportionately high for some profiles. This is owing to the different conditions of profiles collected at very distant locations in the same day. The low-altitude MDM is dominated by two profiles collected in the area of Parma (Italy) for CH_4, while for CO_2, particularly high values in the MDM were noted in four profiles collected in the area of Parma and Krakow (Poland).

After plotting MDM against SRE, two parameters (a single scaling factor and an offset) were derived to train the representation error to match model-data mismatch so that it could be used as a proxy. Values of scaling factors and offset derived using all profiles and a subset of profiles excluding those with high MDM in the mixed layer are shown in Table 5.1.

	Sc. Factor (subset)	Offset (subset)	Sc. Factor (all)	Offset (all)
CO_2	1.95 [ppm/ppm]	0.73 [ppm]	2.94 [ppm/ppm]	1.96 [ppm]
CH_4	3.21 [ppb/ppb]	7.13 [ppb]	10.51 [ppb/ppb]	15.09 [ppb]

Table 5.1: Scaling factor and offset values used to train the representation error derived from all profiles and by considering only a subset without those with unrepresentative values of MDM.

Figure 5.2 ((a) and (c)) shows that by excluding these profiles, the mean MDM is reduced by 52% for CO_2 (from 2.5 ppm to 1.2 ppm) and 54% for CH_4 (from 20.1 ppb to 9.1 ppb). Reductions for the maximum MDM values in proximity of the surface are also important, namely from 7.6 ppm to 5.8 ppm for CO_2, that corresponds to a 23% reduction, and from 68.8 ppb to 21.8 ppb for CH_4, corresponding to a 68% reduction.

In case of a perfect training of SRE using scaling factor and offset, it is expected from the ratios between MDM and ETU to be normally distributed with zero mean and standard deviation equal to one. To verify how good the ETU actually mimics the MDM, the distribution of MDM/SRE values were plotted as box and whiskers for each 1000 m bin (Fig. 5.2 (b) and (d)). The whisker plots show that, apart for some strongly negative outliers in the proximity of the surface, the distribution of the MDM/ETU ratio is quite symmetrical. In addition the single boxes tend to be centered around zero with a standard deviation of one, as it is expected. The overall mean for the MDM/ETU values is -0.21 for CO_2 and 0.10 for CH_4, while the overall standard deviation is 0.97 for CO_2 and 1.01 for CH_4.

By using a subset excluding the profiles with high MDM in the ML, the ETU ranges between 0.8-5.0 ppm for CO_2 SPC and 7.9-24.9 ppb for CH_4 SPC, with mean ETU being respectively 1.3 ppm for CO_2 and 9.5 ppb for CH_4. For vertical partial columns, the ETU is instead about 1.0-4.7 ppm for carbon dioxide and 7.7-28.4 ppb for methane. Mean ETU for CO_2 VPC is 1.9 ppm, while for CH_4 said value is 11.4 ppb (Fig. 5.3, solid lines). Note that for both CO_2 and CH_4, the ETU values for SPC are comparable with the ETU values for VPC. The differences between the expected total uncertainty for VPC and SPC in the lower free troposphere are probably due to slight differences in the location of the lower part of the partial columns when comparing slanted with vertical columns. We regard this as not significant, given the small number of slanted profiles (total of 16) compared to the large number of vertical partial columns (total of 2428).

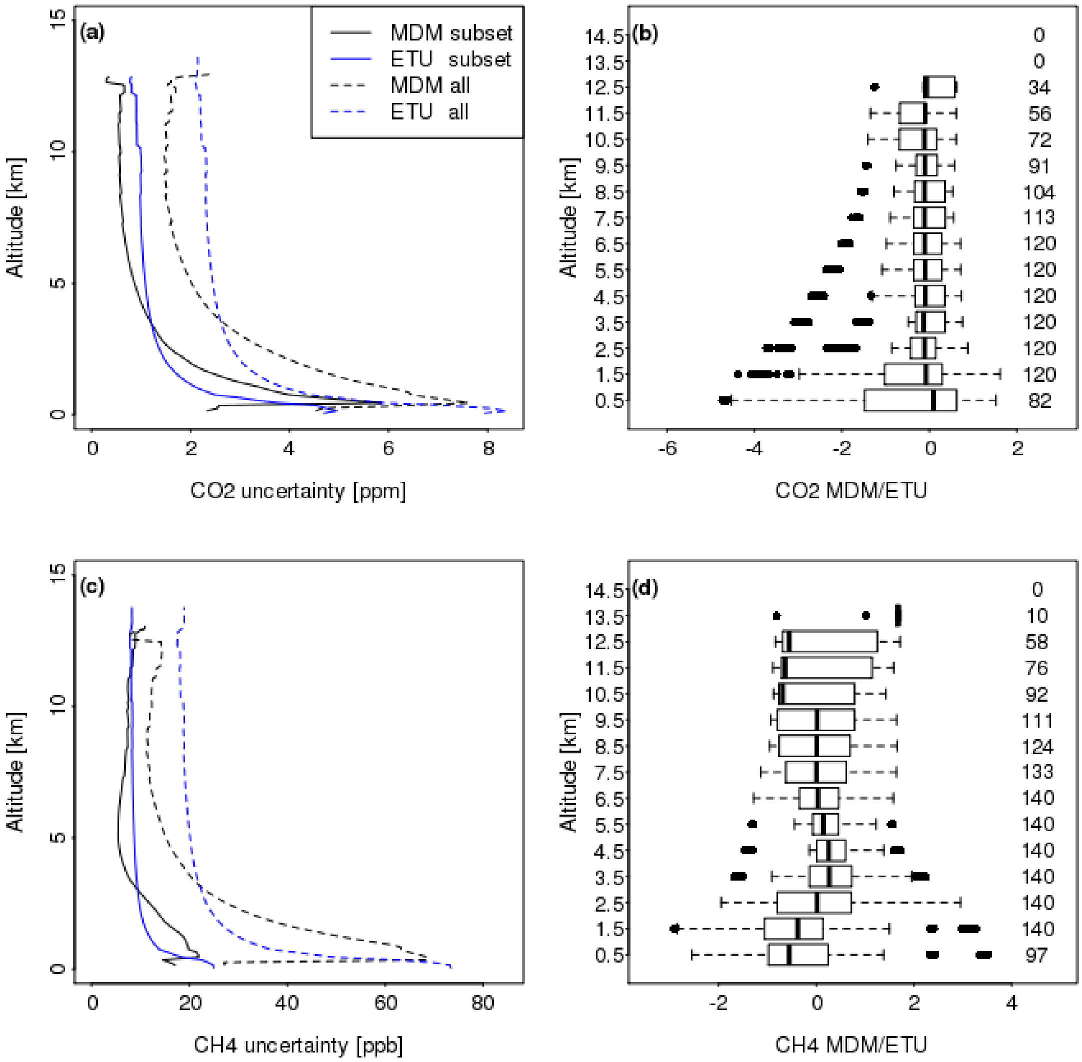

Figure 5.2: Tabs (a) and (c) show values for slanted partial column (SPC) model-data mismatch and expected total uncertainty (ETU) at different heights, for both CO_2 and CH_4. The values of both model-data mismatch and representation error have been derived using all profiles ("all") or by leaving out profiles in Parma for CH_4 and both Parma and Krakow for CO_2 ("subset"). The box-whiskers plots in tabs (b) and (d) show the distribution of individual values of the ratio between model-data mismatch (MDM) and ETU for the "subset" case. The MDM/ETU values are binned every 1000 m and the number of records in each box is reported on the right.

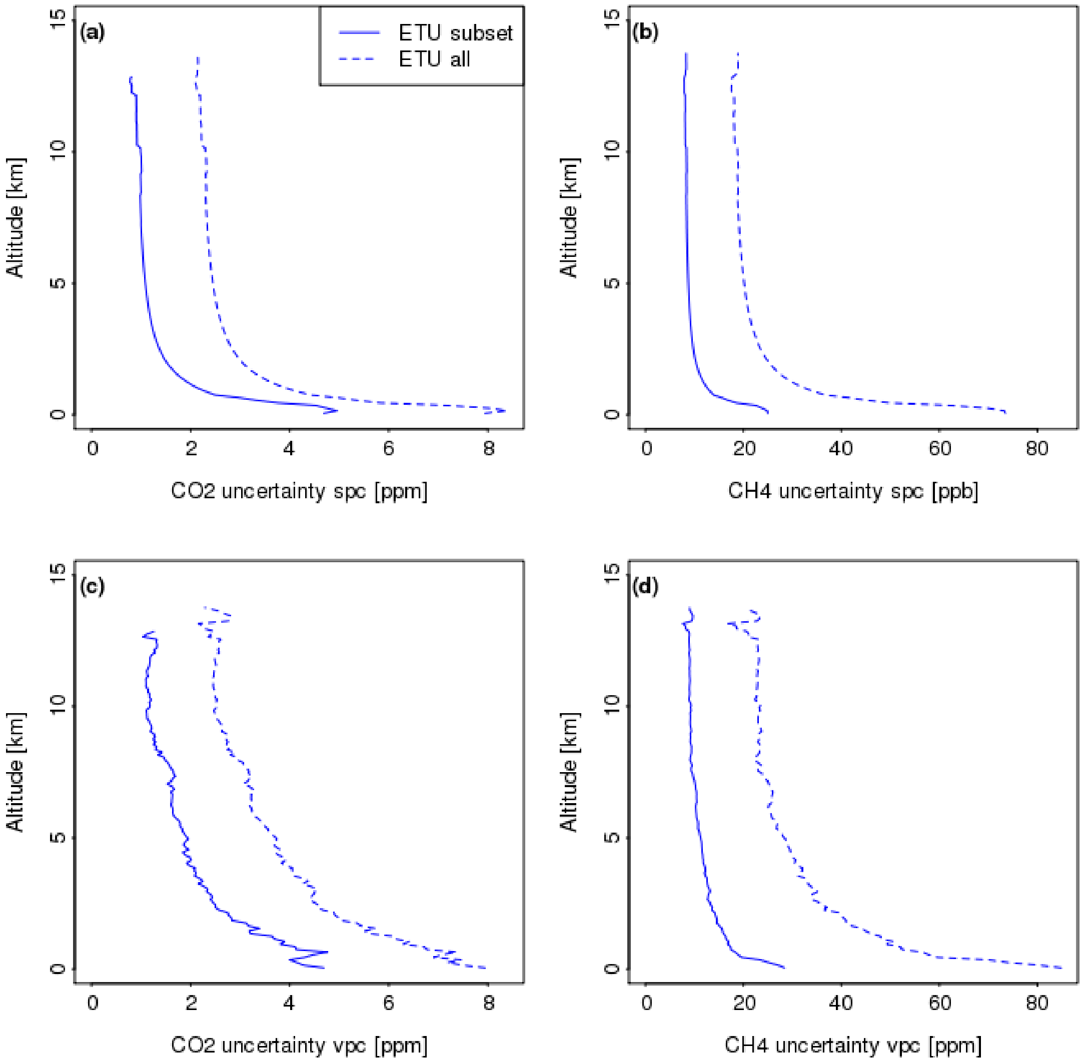

Figure 5.3: Tabs (a) and (c) show representation error values at different heights for CO_2 and CH_4 slanted partial columns. Values have been plotted for all profiles (dashed line) and by removing profiles in Parma and Krakow (solid line), characterized by high spatial variability in the mixed layer. Tabs (b) and (d) shows the same values for vertical partial columns

To investigate the effect of the biosphere on ETU, we run two additional simulations in which only part of the fluxes were modeled. In the first one only anthropogenic emissions were considered, while in the second only biogenic emissions were taken into account. As total CO_2 mixing ratios results from the sum of anthropogenic (“*ant*”) and biogenic (“*bio*”) contributions, the total variance of CO_2 follows Eq. (5.1):

$$var(ant + bio) = var(ant) + var(bio) + 2cov(ant, bio) \qquad (5.1)$$

Figure 5.4 shows the ETU for CO_2 VPC derived by considering only anthropogenic or biogenic fluxes. Interestingly, both of the components show a higher ETU than the total fluxes; this can be achieved only if the two components are negatively correlated (Eq. (5.1)). It was found that the correlation between the two contributions to CO_2 mixing ratio is indeed negative and roughly -0.69. In addition, it is clear from Fig. 5.4b that the high MDM values noted in the mixed layer of profiles collected around Parma and Krakow are almost entirely due to the presence of the biosphere.

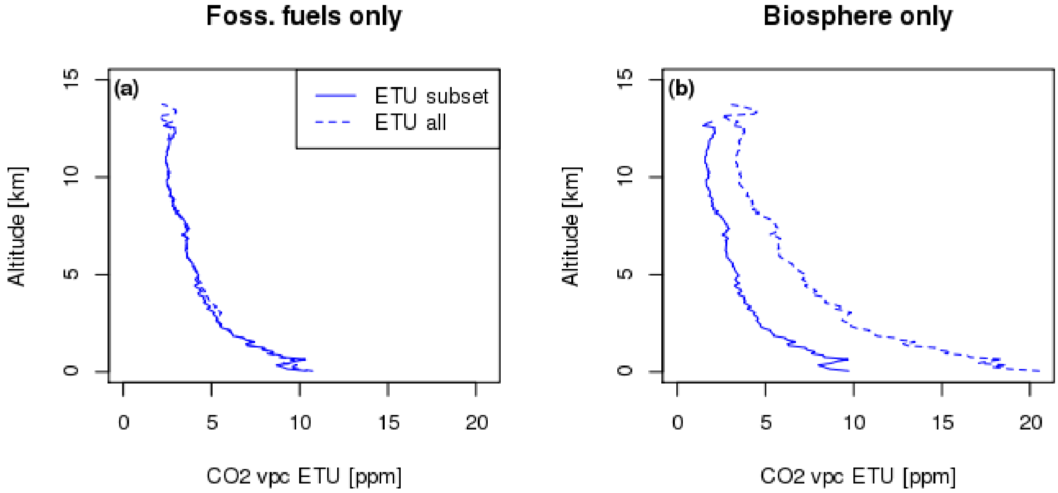

Figure 5.4: Representation error values at different heights for CO_2 vertical partial columns considering only fossil fuels emissions (a) or biospheric emissions (b). Values have been plotted for all profiles (dashed line) and by removing profiles in Parma and Krakow (solid line), characterized by high spatial variability in the mixed layer.

To assess the benefit of using the tracer transport model as a transfer standard to link the column measurements via the calibrated in-situ measurements to the WMO standard scales, we investigate how the relation between the uncertainty in VPC as extrapolated by the transport model and the atmospheric variability of VPC at different times and locations. With no skill of the tracer transport model, the VPC uncertainty would be equal to the atmospheric variability, while a skilled model would be able to explain part of the variability. Figure 5.5 shows the mean VPC values at different heights for both CO_2 and CH_4 together with ETU values and atmospheric variability, here intended as the standard deviation of VPC values from individual profiles. The mean profiles, ETU and atmospheric variability were derived from a subset excluding profiles with strangely large values of MDM, and by considering alternatively the vertical column from the surface upward (indicating a ground return of the lidar signal) or the column starting at 2 km (indicating a typical cloud return of the lidar). VPC is around 400-406 ppm for CO_2 when considering the whole column, and 403-405 ppm when excluding the first 2 km above the ground. For CH_4 such values are instead 1880-1970 ppb and 1875-1895 ppb respectively. For CO_2, atmospheric variability is 0.2-7.2 ppm for the whole vertical domain, and 0.2-4.0 ppm when considering the covered sky case, while for CH4 the atmospheric variability of VPC is about 0.1-26.5 ppb when considering all of the data, and about 2.6-21.9 ppb in the case of cloud cover.

With the exception of CH_4 above around 12 km, the ETU is smaller than the atmospheric variability, namely around 35% for CO_2 and 27% for CH_4. For cloud lidar returns (excluding the lowest 2 km), the ETU is roughly 50% lower than atmospheric variability for CO_2, and 44% for CH_4. The values for the ETU indicate the uncertainty in traceability of lidar measurements back to WMO standard scales; for CO_2, the mean ETU is 1.9 ppm for ground returns and 0.9 ppm for cloud returns, while for CH_4 these values are 11.4 ppb and 7.8 ppb respectively.

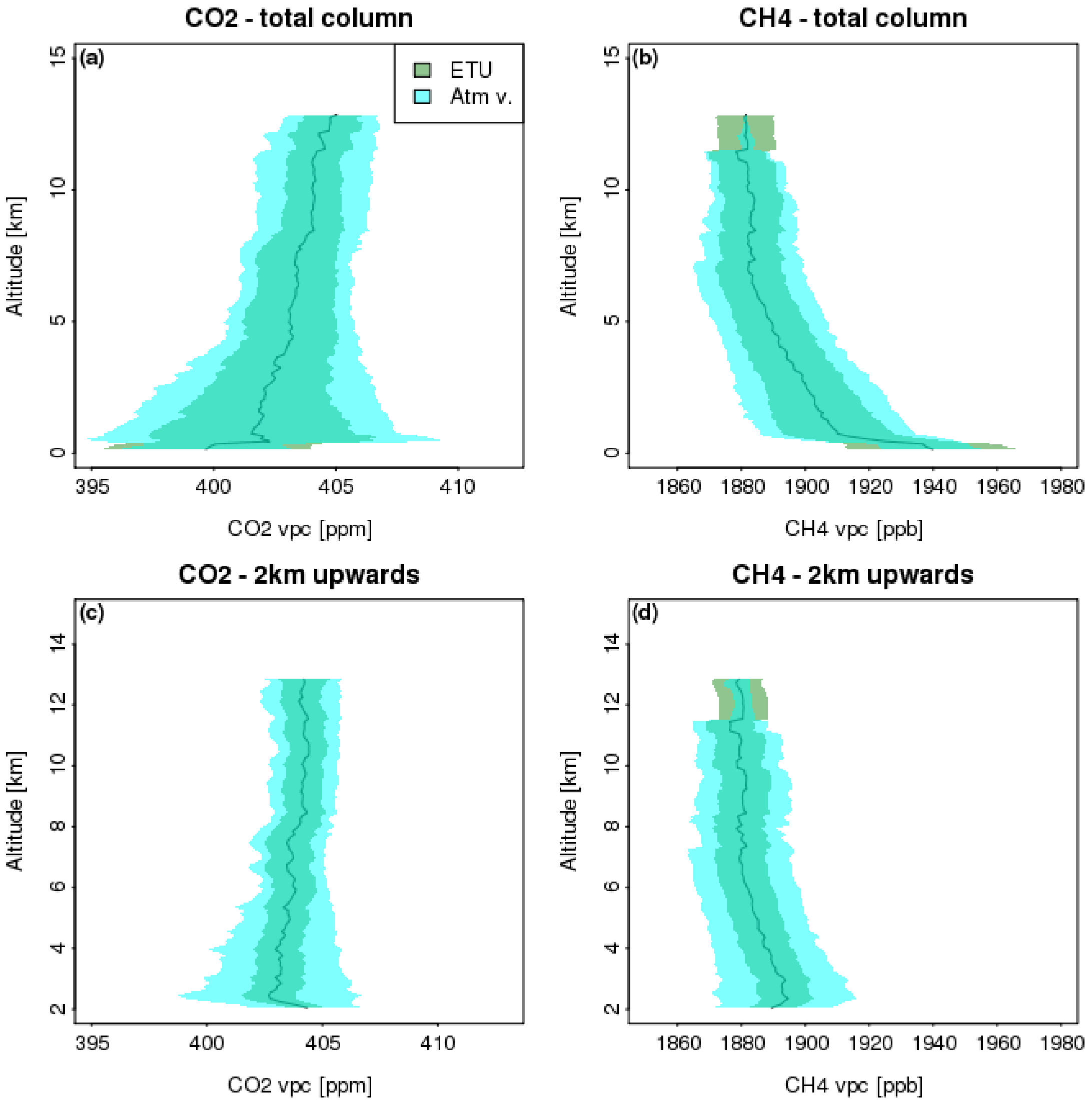

Figure 5.5: Mean VPC profile for CO_2 and CH_4 derived by excluding the high-MDM ones considering the whole vertical column or by excluding the lower 2 km. The values of ETU and atmospheric variability (sd of individual VPC profiles) are plotted in green and cyan respectively.

Finally, for linking CHARM-F to global WMO standards, the observed lidar retrievals will have to be adjusted to match the simulated columns (calibrated by the in-situ observations). Ideally this can be achieved using a single calibration factor as done e.g. for TCCON (the Total Carbon Column Observing Network) a network of ground based remote sensing systems (Wunch et al., 2010). Using all profiles collected during the CHARM-F test campaign, the resulting uncertainty of such a species-specific scaling factor would thus be smaller than those shown in Fig. 5.5 by a factor equal to the square root of the number of available profiles, namely 12 for CO_2 and 14 for CH_4. This results in an overall uncertainty in the traceability of 0.5 ppm for CO_2 and 3 ppb for CH_4. Note that the corresponding WMO compatibility targets are 0.1 ppm and 2.0 ppb respectively, which shows that further improvements are needed.

Chapter 6: Summary and Outlook

As already explained in the introduction, this thesis is divided into three main parts for a corresponding number of topics. In the first part we describe different sources of error in the modeling of in-situ airborne measurements from IAGOS, with particular emphasis on the representation error. Building on this first part, in the second part we derive a method to improve the constraint of multi-species inversion in a computationally cheap way. This shows the potential for future carbon budget studies based on future IAGOS measurements of CO_2, CO and CH_4. Finally, the third part is focused on validation of the CHARM-F airborne lidar. Such instrument is part of the HALO aircraft payload and is not yet traceable to the WMO standards. By using in-situ measurements from the same platform and a modeling framework, we derived a method to quantify the traceability to WMO calibration scales of CHARM-F vertical partial columns for future measurement campaigns.

6.1 Sources of uncertainty in modeling airborne measurements: conclusions

In Chapter 3 we quantitatively described the contribution of the three major model-derived uncertainty sources: mismatch in the mixed layer depth (related to uncertainty in the transport model), bias in the fluxes provided by the emission inventories, and spatial representation error. We have shown the contribution for both random and bias component of the different model-derived error categories for both STILT/EDGAR simulations and MACC reanalysis. Both models show similar contributions for the random component of mixed layer depth and flux accuracy with the former clearly dominating on the latter. The bias component for mixed layer depth and flux accuracy has opposite sign in both models. In addition, is comparable in magnitude for STILT/EDGAR whereas in MACC the flux error is clearly more important. Contributions from spatial representation error tend to increase with degraded resolutions.

The clear dependence of representation error on both spatial resolution and wind direction shows that observations are influenced by local emissions; so high-resolution emission inventories are to be preferred for modeling applications such as inverse modeling. It is however worth to point out that higher resolution models will have a direct impact on the computational effort and hence temporal and financial resources required to carry out such simulations. However, in this paper we have shown that spatial and temporal variation in the representation error allows for describing about one sixth to one fifth of the variance in model-data mismatch. Therefore, it is likely that information on temporal and spatial variation of representation error derived from the STILT/EDGAR simulations can be used to improve quantitative analyses when using coarse models. On the other hand, given that the spatial representation error, though significant, does not dominate the model-data mismatch, MOZAIC/IAGOS profile data can be regarded as spatially representative to a certain degree. This is good news for such a dataset, which has been collected at many locations around the globe in the vicinity of major metropolitan areas.

When considering the whole dataset, uncertainty in simulated vertical transport showed that a low bias in mixed layer depth of 43% for EDGAR/STILT and 37% for MACC results in a reduction of CO enhancements of 6.9% and 0.3% respectively in the most representative location. We also provided evidence that the modeled outputs for CO underestimate observations by a factor 2-2.5, suggesting a bias in the emission inventories. As most of the CO enhancements in the mixed layer is caused by regional fluxes, rather than advected contributions from the lateral boundary, the difference between observed and modeled enhancements can be seen as the difference between actual and estimated emissions. Observations are needed to correct this underestimation, as current emission inventories are likely to perform poorly, e.g. in forecasting of regional air pollution. Flux error is targeted in modeling to provide top-down constraint on emission inventories; however the accuracy in top-down constrained emissions is affected by the other sources of model-data mismatch error. For example, uncertainties in mixed layer heights and spatial representation error limit the accuracy to 30-60% (bias components in Fig. 3.16), thus they need to be taken into account in order to provide a more accurate top-down constraint. Note that uncertainties introduced by the (simplified) representation of photochemistry and by the choice of lateral boundary condition are rather small.

Airborne measurements are not often used for inverse (top-down) modeling of fluxes. However, given that observations from commercial airliners provide vertical profile information near major cities around the globe, and given that the emissions can be constraint by the enhancement within the boundary layer as shown here, we argue that they should be considered in inverse modeling. The method of error partitioning described in this paper will be especially important in the context of the upcoming availability of CO_2 and CH_4 profile data within IAGOS. In fact, the availability of three carbon-based tracer gases (CO, CO_2 and CH_4) will allow for multi-species inversion studies making use of a transport-flux coupling as the one described here. In this context, with different tracers that share the same transport, and that share part of their emission categories, the ability to discriminate between different sources of uncertainty will be useful. Also the use of CO as a proxy for the anthropogenic emission contribution to the CO_2 mixing ratio will be improved with the better described contributions of uncertainties to model-data mismatch. The analysis performed in Chapter 4 tries to pave the way to this kind of studies.

Furthermore, as MOZAIC/IAGOS is also an important data provider for validation of satellite observations (Cal/Val), a possible future study is to perform a similar analysis aiming to assess the contributions of different sources of uncertainty contributing to satellite-airborne mismatches. A feasible option for the satellite data are the measurements from the MOPITT sensor onboard the NASA’s Earth Observing System Terra spacecraft. A further alternative is represented by the IASI sensor onboard the ESA’s MetOp satellite. The implementation into the

IAGOS database of the information on part of the error partitioning presented in this study is envisioned in the frame of the IGAS project as first application of our results.

6.2 Multi-species inversions: conclusions

Building upon the work described in Chapter 3, Chapter 4 presents a synthetic experiment aiming to evaluate the effects of exploiting correlations between different trace gases in an atmospheric inversion. We quantitatively described the capability of the modeling framework to reproduce observations, the performance of the inversion scheme in reducing the uncertainty of the different trace gases, and the benefits of multi-species inversions compared to corresponding single-species inversions. We also describe a method to re-scale different prior uncertainty covariance matrices so that the corresponding posterior uncertainties are actually comparable.

Where possible, we confronted model outputs with available observations. Such comparison, possible only for CO, showed a good degree of agreement between the model and observations with an overall correlation of roughly 0.75; modeled values for CO enhancement underestimate the observed ones by a factor of roughly 3.8, compatible with what was found in Chapter 3 for the year 2011. It is found that posterior uncertainty is much lower than the prior for all of the five simulated species. The mean uncertainty reduction for CO_2 emissions from fossil fuels is roughly 38%, for GEE it is around 41% while for respiration it is roughly 44%. For CO and CH_4 the uncertainty reduction is about 63% and 67% respectively. Finally, we described quantitatively the benefit of using multi-species inversions by exploiting correlations in different chemical species. It is found that considering correlations between different trace gases can reduce the posterior uncertainty by up to about 20% for monthly fluxes. These benefits are however dependent on the error structure of the prior uncertainty.

The present paper paves the way for future studies using simultaneous measurements of different trace gases. This will be especially important in the context of the upcoming routine measurements of CO_2, CO, and CH_4 vertical profiles within IAGOS. As IAGOS makes use of commercial airliners, such profiles will be collected in the vicinity of major international airports, and hence in the vicinity of major metropolitan areas, where many different human activities take place simultaneously. In such a context, any improvement in the constraint of atmospheric inversions will be particularly useful. A possible improvement in this analysis would be to evaluate the effects of different correlation factors specific to different pairs of anthropogenic species, fuels and emission sectors.

6.3 Lidar validation: conclusions

Chapter 5 makes an attempt to quantitatively link measurements made by the CHARM-F lidar carried by the HALO aircraft to WMO standards for CO_2 and CH_4 at a time when observations from CHARM-F are still not ready for direct comparison. To achieve this goal we use in-situ observations from the JIG sensor, also carried by HALO, and a modelling framework centred on the STILT transport model. The STILT modelling framework was calibrated to match in-situ observations and then used to estimate the uncertainty of future vertical columns from the CHARM-F lidar and make it traceable to WMO atmospheric standard.

We describe a correction model to correct in-situ profiles that makes use of three linear fits for different atmospheric regimes: mixing layer, free troposphere and stratosphere. After calibration of the model, the mean model-data mismatch (MDM) for CO_2 is reduced by 91% from -1.1 ppm to -0.1 ppm, and the standard deviation of the MDM goes from 2.2 pm to 1.8 ppm, corresponding to an 18% reduction. Similar results are found for CH_4, as mean MDM is reduced from 25.3 ppb to -0.4 ppb (98%), while the standard deviation of the model-data mismatch goes from 27.2 ppb to 21.7 ppb (20 % reduction). It is found that profiles collected in the area of Parma (Italy) and Krakow (Poland) have strangely high values of MDM in the mixing layer. For CO_2, it was established that this is largely due to the variability of the biogenic simulated fluxes.

After bias-correcting profiles with the above-mentioned model, slanted partial columns were derived from in-situ observations and model output along the flight track. Representation error (SRE) of slanted partial column was then trained to match the overall MDM and the corresponding SRE for vertical partial column was derived in a similar fashion. For this reason, the trained SRE (ETU, Expected Total Uncertainty) for vertical partial column is here assumed to be the expected uncertainty of future observed retrievals from CHARM-F. ETU for slanted partial columns ranges between 0.8-5.0 ppm for CO_2 and 7.9-24.9 ppb for CH_4, with mean ETU being respectively 1.3 ppm for CO_2 and 9.5 ppb for CH_4. For vertical partial columns, ETU is instead about 1.0-4.7 ppm for carbon dioxide and 7.7-28.4 ppb for methane. Mean representation error for CO_2 vertical partial column is 1.9 ppm, while for CH_4 said value is 11.4 ppb.

Values of ETU for VPC were finally compared with mean values of modeled VPC and atmospheric variability arising from differences of individual VPC profiles. Such test was performed considering the whole vertical column below the flight track, or by excluding the first 2 km above the ground, in an effort to simulate the conditions associated with a lidar return by clouds. It is found that ETU is between 27% and 50% lower than atmospheric variability, indicating that using a transport model as a transfer standard does indeed reduce the uncertainty when comparing remote sensing column observations to in-situ measurements. For CO_2, the mean ETU is 1.9 ppm for ground returns and 0.9 ppm for cloud returns, while for CH_4 these values are 11.4 ppb and 7.8 ppb respectively. Such values for ETU indicate the uncertainty in traceability of lidar measurements back to WMO standard scales.

The validation of CHARM-F marks an important milestone not only for the HALO aircraft, but also for the MERLIN space mission from ESA, whose lidar is expected to build upon the experience with CHARM-F. The obvious next step once observed partial columns will be ready for users is to compare the calibrated STILT vertical partial column with CHARM-F retrievals. In addition, as WMO traceability target for surface measurements are 0.1 ppm CO_2 and 2 ppb CH_4, a further goal for future analysis will be the improvement of the modelling framework to match such targets. One contribution would be the improvement of the modeling system by using optimized fluxes and a transport model with higher resolution (e.g. STILT-WRF, Pillai et al, 2012). Further improvements can be achieved by selecting specific flight patterns minimizing the spatial mismatch between lidar and in-situ (e.g. spiral ascent or descent) during future campaigns.

Bibliography

Abshire, J.B., Riris, H., Allan, G.R., Weaver, C.J., Mao, J., Sun, X., Hasselbrack, W.E., Kawa, S.R. and Biraud, S., 2010. Pulsed airborne lidar measurements of atmospheric CO2 column absorption. *Tellus B*, ***62*****(5)**, pp.770-783.

Bergamaschi, P., Hein, R., Heimann, M., Crutzen, P. J., 2000. Inverse modeling of the global CO cycle: 1. Inversion of CO mixing ratios. *J. Geophys. Res.*, **105(D2)**.

Bergamaschi, P., co-authors, 2009. Inverse modeling of global and regional CH4 emissions using SCIAMACHY satellite retrievals. *J. Geophys. Res.*, **114(D22)**.

Brioude, J., co-authors, 2013. Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: assessing anthropogenic emissions of CO, NO_x and CO_2 and their impacts. *Atmospheric Chemistry and Physics*, ***13*****(7)**, 3661–3677.

Boschetti, F., Chen, H., Thouret, V., Nedelec, P., Janssens-Maenhout, G., Gerbig, C.: On the representation of IAGOS/MOZAIC vertical profiles in chemical transport models: contribution of different error sources in the example of carbon monoxide, Tellus B, **67**, 2015.

Bousquet, P., co-authors, 2006. Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature*, **443**, 439-443.

BP (British Petroleum), 2014. Statistical Review of World Energy: http: //www.bp.com/statisticalreview, last access: September 2014.

Browell, E.V., Dobbs, M.E., Dobler, J., Kooi, S., Choi, Y., Harrison, F.W., Moore III, B. and Zaccheo, T.S., 2008, June. Airborne demonstration of 1.57-micron laser absorption spectrometer for atmospheric CO2 measurements. In*24th International Laser Radar Conference, Boulder, Colorado* (pp. 23-24), 2008

Camia, A., co-authors, 2009. European Commission, Joint Research Centre (JRC) - Forest fires in Europe 2008. *Office for Official Publications oft he European Communities*.

Denier van der Gon, H. D., Hendriks, C., Kuenen, J., Segers, A., and Visschedijk, A.: Description of current temporal emission patterns and sensitivity of predicted AQ for temporal emission patterns, TNO Report, EU FP7 MACC deliverable report D_D- EMIS_1.3, available at: https://gmes-atmosphere.eu/documents/ deliverables/d-emis/MACC_TNO_del_1_3_v2.pdf, 2011.

Ehret, G., Flamant, P., Alpers, M., Millet, B., Stephan, C. and Crebassoi, P., 2013, June. The French–German climate mission MERLIN. In Coherent Laser Radar Conference (CLRC 2013), Barcelona, Spain.

Elguindi, N., co-authors, 2010. Current status of the ability of the GEMS/MACC models to reproduce the tropospheric CO vertical distribution as measured by MOZAIC. *Geoscientific Model Development*, **3**, 501-518.

European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL), 2011. Emission Database for Global Atmospheric Research (EDGAR), release version 4.1 and 4.2. http://edgar.jrc.ec.europa.eu.

Filges, A., Gerbig, C., Chen, H., Franke, H., Klaus, C. and Jordan, A.: The IAGOS-core greenhouse gas package: a measurement system for continuous airborne observations of CO 2, CH 4, H 2 O and CO, Tellus B, **67**, 2015.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I. and Bala, G.: Climate–carbon cycle feedback analysis: results from the C4MIP model intercomparison, Journal of Climate, **19(14)**, pp.3337-3353, 2006.

Gerbig, C., co-authors. Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms. *J Geophys Res-Atmos*, **108(D24)**, 2003a.

Gerbig, C., co-authors. Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. *J Geophys Res-Atmos*, **108(D24)**, 2003b.

Gerbig, C., Körner, S. and Lin, J.C.. Vertical mixing in atmospheric tracer transport models: error characterization and propagation. *Atmospheric Chemistry and Physics*, ***8*(3)**, pp.591-602, 2008

Gettelman, A., M. L. Salby, and F. Sassi. "Distribution and influence of convection in the tropical tropopause region." *Journal of Geophysical Research: Atmospheres* **107(D10)**, 2002.

Gourdji, S. M., co-authors. North American CO_2 exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion. *Biogeosciences*, ***9*(1)**, 457–475, 2012

Heimann, M., and Körner, S.: The global atmospheric tracer model TM3. Technical Reports, Max-Planck-Institut für Biogeochemie, **5**, pp.131, 2003.

Hein, R., Crutzen, P. J., and Heimann, M: An inverse modeling approach to investigate the global atmospheric methane cycle. *Global Biogeochemical Cycles*, **11(1)**, 43-76, 1997.

Inness, A., co-authors, 2013. The MACC reanalysis: An 8 yr data set of atmospheric composition. *Atmospheric Chemistry and Physics*, **13**, 4073-4109.

IEA (International Energy Agency) World Energy Outlook 2014, http://www.iea.org/bookshop/720-World_Energy_Outlook_2014, 2014.

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II, and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)], IPCC, Geneva, Switzerland, 151 pp., 2014

Janssens-Maenhout, G., Pagliari, V., Guizzardi, D., Muntean, M., 2013. Global emission inventories in the Emission Database for Global Atmospheric Research (EDGAR) – Manual (I): Gridding: EDGAR emissions distribution on global gridmaps. JRC Report, EUR 25785 EN.

Jung, M., Henkel, K., Herold, M. and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, Remote Sensing of Environment, **101(4)**, 534-553, 2006.

Kadygrov, N., Broquet, G., Chevallier, F., Rivier, L., Gerbig, C. and Ciais, P.: On the potential of the ICOS atmospheric CO 2 measurement network for estimating the biogenic CO 2 budget of Europe. Atmos. Chem. Phys., **15(22)**, 12765-12787, 2015.

Kopacz, M., Jacob, D. J., Henze, D. K., Heald, C. L., Streets, D. G. Zhang, Q., 2009. Comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of carbon monoxide using satellite (MOPITT) measurements of CO columns. *J Geophys. Res*., **114(D4)**.

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, T. F. and Heimann, M.: Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part2: the European terrestrial CO_2 fluxes, Atmos. Chem. Phys. Discuss., **1–44**, doi:10.5194/acp-2016-578, 2016.

Kretschmer, R., Gerbig, C., Karstens, U. and Koch, F. T., 2012. Error characterization of CO2 vertical mixing in the atmospheric transport model WRF-VPRM. *Atmos. Chem. Phys*., **12(5)**, 2441–2458.

Lin, J., co-authors, 2003. A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *J. Geophys. Res*., **108(D16)**, 2003.

Lin, J., co-authors, 2004. An empirical analysis of the spatial variability of atmospheric CO 2: Implications for inverse analyses and space-borne sensors. *Geophys. Res. Lett*., **31(23)**, 2004.

Lin, J.,Gerbig, C., 2005. Accounting for the effect of transport errors on tracers inversions. *Geophys. Res. Lett*., **32(1)**, 2005.

Machida, T., Matsueda, H., Sawa, Y., Nakagawa, Y., Hirotani, K., Kondo, N., Goto, K., Nakazawa, T., Ishikawa, K. and Ogawa, T., 2008. Worldwide Measurements of Atmospheric CO 2and Other Trace Gas Species Using Commercial Airlines, *J. Atmos. Oceanic Technol*., **25(10)**, 1744–1754, doi:10.1175/2008JTECHA1082.1.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C.: A satellite-based biosphere parameterization for net ecosystem CO 2exchange: Vegetation Photosynthesis and Respiration Model (VPRM). Global Biogeochemical Cycles, **22(2)**, n/a–n/a. doi:10.1029/2006GB002735, 2008.

Marenco, A., Thouret, V., Nédélec, P., Smit, H., Helten, M., Kley, D., Karcher, F., Simon, P., Law, K. and Pyle, J., 1998. Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, An overview, *Journal of Geophysical Research: Atmospheres* (1984–2012), **103(D19)**, 25631–25642.

Nassar, R., co-authors, 2011, Inverse modeling of CO_2 sources and sinks using satellite observations of CO2 from TES and surface flask measurements. *Atmospheric Chemistry and Physics*, ***11*(12)**, 6029-6047.

Nedelec, P., co-authors, 2003. An improved infrared carbon monoxide analyser for routine measurements aboard commercial airbus aircraft: Technical validation and first scientific results of the MOZAIC III program. *Atmospheric Chemistry and Physics Discussions*, **3(4)**, 3713-3744.

Nedelec, P., Thouret, V., Brioude, J., Sauvage, B., Cammas, J. P., Stohl, A., 2005. Extreme CO concentrations in the upper troposphere over northeast Asia in June 2003 from the in situ MOZAIC aircraft data. *Geophysical Research Letters*, **32(14)**.

Palmer, P. I., Suntharalingam, P., Jones, D. B. A., Jacob, D. J., Streets, D. G., Fu, Q.: Using CO2:CO correlations to improve inverse analyses of carbon fluxes. Journal of Geophysical Research, **111(D12)**, D12318. doi:10.1029/2005JD006697, 2006.

Petzold, A., Thouret, V., Gerbig, C., Zahn, A., Brenninkmeijer, C.A.M., Gallagher, M., Hermann, M., Pontaud, M., Ziereis, H., Boulanger, D., Nédélec, P., Smit, H.G.J., Cammas , J.-P., Volz-Thomas, A., and the IAGOS Team, 2015, Global-Scale Atmosphere Monitoring by In-Service Aircraft – Current Achievements and Future Prospects of the European Research Infrastructure IAGOS, *Tellus B*, this issue.

Pierangelo, C., B. Millet, F. Esteve, M. Alpers, G. Ehret, P. Flamant, S. Berthier et al. "MERLIN (methane remote sensing LIDAR mission): an overview." In *EPJ Web of Conferences*, vol. 119, p. 26001. EDP Sciences, 2016.

Pillai, D., Gerbig, C., Kretschmer, R., Beck, V., Karstens, U., Neininger, B. and HEIMANN, M.: Comparing Lagrangian and Eulerian models for CO_2 transport – a step towards Bayesian inverse modeling using WRF/STILT-VPRM, Atmos. Chem. Phys., **12(19)**, 8979–8991, doi:10.5194/acp-12-8979-2012, 2012.

Protonatariou, A., Tombrou, M., Giannakopulos, C.,Kostopoulou, E.,Le Sager, P., 2010. Study of CO surface pollution in Europe based on observations and nested-grid applications of GEOS-CHEM global chemical transport model. *Tellus B*, **62(4)**, 209-227.

Riris H., K. Numata, S. Li, S. Wu, A. Ramanathan, M. Dawsey, J. Mao, R. Kawa, and J. B. Abshire. "Airborne measurements of atmospheric methane column abundance using a pulsed integrated-path differential absorption lidar." Applied Optics 51, no. 34, 8296-8305, 2012.

Rödenbeck, C., Houweling, S., Gloor, M. and Heimann, M.,2003. Time-dependent atmospheric CO2 inversions based on interannually varying tracer transport. *Tellus B*, **55(2)**, pp. 488–497.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing Co. Pte. Ltd., Singapore, 238 pp., 2000.

Sasakawa, M., Shimoyama, K., Machida, T., Tsuda, N., Suto, H., Arshinov, M., Davydov, D., Fofonov, A., Krasnov, O., Saeki, T. and Koyama, Y.: Continuous measurements of methane from a tower network over Siberia.Tellus B, **62(5)**, pp.403-416, 2010.

Seibert, P., F. Beyrich, S. E. Gryning, S. Joffre, A. Rasmussen, and P. Tercier: Review and intercomparison of operational methods for the determination of the mixing height. *Atmos. Environm.,* **34(7)**, 1001-1027, 2000.

Spiers, G.D., Menzies, R.T., Jacob, J., Christensen, L.E., Phillips, M.W., Choi, Y. and Browell, E.V.: Atmospheric CO 2 measurements with a 2 μm airborne laser absorption spectrometer employing coherent detection.*Applied optics*, ***50*****(14)**, pp.2098-2111, 2011.

Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C. and Mukai, H., 2011. The CO2 release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios. *Atmos. Chem. Phys.*, **11(14)**, 6855–6870.

Stein, O., co-authors, 2014. On the wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations. *Atmos. Chem. Phys*., **14(17)**, 9295–9316.

Streets, D.G., Bond, T.C., Carmichael, G.R., Fernandes, S.D., Fu, Q., He, D., Klimont, Z., Nelson, S.M., Tsai, N.Y., Wang, M.Q. and Woo, J.H.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, Journal of Geophysical Research: Atmospheres, **108(D21)**, 2003.

Vardag, S. N., Gerbig, C., Janssens-Maenhout, G., and Levin, I.: Estimation of continuous anthropogenic CO_2: model-based evaluation of CO_2, CO, $\delta^{13}C(CO_2)$ and $\Delta^{14}C(CO_2)$ tracer methods, Atmos. Chem. Phys., **15**,, doi:10.5194/acp-15-12705-2015, 2015.

Volz-Thomas, A., Cammas, J.-P., Brenninkmeijer, C. A., Machida, T., Cooper, O. R., Sweeney, C. and Waibel, A., 2009. Civil Aviation Monitors Air Quality and Climate, *EM Magazine, Air & Waste Management Association*, p. 16-19, October 2009.

Wang, H., Jacob, D.J., Kopacz, M., Jones, D.B.A., Suntharalingam, P., Fisher, J.A., Nassar, R., Pawson, S. and Nielsen, J.E.: Error correlation between CO 2 and CO as constraint for CO 2 flux inversions using satellite data, Atmospheric Chemistry and Physics, **9(19)**, pp.7313-7323, 2009.

Wunch, D., Toon, G.C., Wennberg, P.O., Wofsy, S.C., Stephens, R.S., Fischer, M.K., Uchino, O., Abshire, J., Bernath, P., Biraud, S.C. and Blavier, J.F.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, 2010

Zbinden, R. M., Thouret, V., Ricaud, P., Carminati, F., Cammas, J. P., and Nédélec, P., 2013. Climatology of pure tropospheric profiles and column contents of ozone and carbon monoxide using MOZAIC in the mid-northern latitudes (24° N to 50° N) from 1994 to 2009. *Atmos. Chem. Phys.*, **13**, 12363-12388.

Acknowledgements

First and foremost I'd like do express my deepest gratitude to my "principal" advisor Dr. Habil. Christoph Gerbig for his constant support and guidance. I have been extremely lucky to have such a knowledgeable, patient and motivating supervisor who cared so much about my work and was available to help whenever needed. His positive attitude and constructive criticism played an extremely important role in shaping this dissertation. He was the best advisor I could possibly imagine to have for my PhD.

I also wish to thank my supervisor at the university, Prof. Dr. Kai Uwe Totsche for his helpful comments during PAC meetings during these four years.

I am much obliged to Dr. Julia Marshall for her kind support at the beginning of the PhD, during the absence of Christoph, and for giving insightful comments on papers, during PAC meetings and for teaching me the basis of how to handle satellite data.

Many thanks to Dr. Thomas Koch for teaching me the very first notions of linux and modeling in general, and to all of the other members of the department for useful suggestions about the project and fruitful discussions covering the most different topics.

On a more logistical note, I am indebted to Steffi Rothhardt, John Schmidt, Anna Görner and Ulrike Schleier for their kind help in settling down in a new country and for assisting me with many administrative issues throughout these last years.

As the whole work for the thesis was done on the cluster of the institute, I also would like to thank the IT department for their professionalism and constant effort in maintaining the HPC cluster.

A special thank to my family whose unconditional love and support throughout my whole life gave me the opportunity to eventually accomplish this important milestone. Thank you for everything, I could have not done anything without you.

Last but not the least, many thanks to a lot of kind and nice people that made my life in Jena much enjoyable. You are too many to be listed, so I will not even try it, for fear of forgetting somebody. I just want you to know that these have been the happiest years in my life so far, and I owe most of it to you. Thanks.

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und unter Verwendung der angegebenen Hilfsmittel, persönlichen Mitteilungen und Quellen angefertigt habe.

Jena, Fabio Boschetti

Curriculum Vitae

Personal information	
First name:	Fabio
Family name:	Boschetti
Date of birth:	Sep 25th, 1986
Place of birth:	Desenzano del Garda, Italy
Email:	fabosk@hotmail.it
Education	
06/2013 – 06/2017	• PhD researcher at Max Plank Institute for Biogeochemistry (Jena, Germany)
09/2007 – 12/2012	• Master in Physics at Universita' Cattolica del Sacro Cuore (Brescia, Italy)
Experience and Skills	
2013 – 2017	Researcher, Max Plank Institute for Biogeochemistry
Communication:	• Strong communication/presentation skills: presented data clearly in both internal seminars and international conferences • Adapting style and content of the presentation according to the audience • Strong synthetizing and writing skills. Wrote scientific articles in international peer-reviewed journals. Wrote regular progression reports. Communicate results also to non-experts
Soft skills:	• Communicated and collaborated at all professional levels, with people from diverse countries/cultures. Able to work as team player or independent • Experienced in analyzing large amounts of data, analytical thinking to identify and solve problems • Managed projects and collaborations. Used to work in advance and respect deadlines

Jena, Fabio Boschetti